On Deliberative Authoritarian Governance*

Jidong Chen† Yiqing Xu‡

(job market paper)

Abstract

Why would an authoritarian regime set up deliberative institutions to allow people to complain publicly if, as is often presumed, complaints facilitate protests and cause social instability? We argue that deliberation is a process of hierarchical communication not only between the government and the citizens, but also among the citizens. We show that deliberation serves two functions. First, it helps the government respond to fluctuating public opinion. Second, it can also help to reshape the citizens' beliefs. Specifically, deliberation disorganizes citizens if they find themselves split over government policies. However, if deliberation reveals that a protest can be successful, the government identifies the danger, and improves the policy to appease the opposition. When the citizens are perceived to be sufficiently homogeneous, deliberation is allowed. We further investigate two deliberative mechanisms that combine a private poll with either a committed responsiveness or a censorship strategy. To achieve the best possible payoff in both mechanisms, the government needs to constrain itself from over-manipulating citizens' learning, thus amplifying the disorganization effect.

KEYWORDS: authoritarian rule, hierarchical communication, deliberation, collective action

---

*This version: August 29, 2013. We are indebted to Carles Boix, Bruce Bueno de Mesquita, Bob Gibbons, Jens Hainmueller, Haifeng Huang, Matias Iaryczower, John Londregan, Nolan McCarty, Adam Meirowitz, Kris Ramsay, Tom Romer, Peter Rosendoff, Alastair Smith, Jim Sndyer, Lily Tsai, Leonard Wantchekon, as well as seminar participants at Princeton, MIT, Harvard-MIT-BU China Workshop, New Faces in China Studies at Duke, Political Economy Graduate Conference at NYU and Biennial Conference of China Development Studies at Shanghai Jiao Tong University, for helpful comments. We are also grateful for Q-APS program for generous support.

†Department of Politics and Program in Political Economy, Princeton University. Email: jidongc@princeton.edu.

‡Department of Political Science, Massachusetts Institute of Technology. Email: xyq@mit.edu.

To silence the populace is as grim a task as preventing flood. A blocked river would eventually inundate and cause great catastrophe; the same can be said of a stifled people. It is therefore wise to dredge the river to let it run free, and to enable the people to speak its mind.

——Discourses of the States (*Guo Yu*), around 500 BC

1 Introduction

In order to maintain social stability and stay in power, an authoritarian incumbent has to find to a way to please and/or repress citizens under its rule (Svolik, 2012). One popular argument is that sophisticated authoritarian rulers provide citizens with economic benefits, and prevent them from coordinating with each other in the political domain (Bueno de Mesquita and Downs, 2005). Political repression is no doubt a feature of authoritarian rule, but an authoritarian state like China also shows a relatively high degree of openness and responsiveness in terms of allowing people to speak and meeting their voiced needs. A recent survey shows that around 10% of urban dwellers in large Chinese cities complain about the local government or government officials on daily matters through government-provided channels, much more than those who contact the media or initiate collective action such as a protest (Table 1 and Figure 1).[1] Among them, 54.8% report that complaining helps solve their problems. Data also show that, apart from reporting officials' misconduct, more often citizens use complaints to demand better public services

[1]The data come from "China Public Governance Survey" (CPGS), a research project carried out by Unirule Institute of Economics and HorizonKey. The authors recognize their assistance in providing the data. However, the views expressed herein are the authors' own and do not necessarily represent the views of the CPGS project. In 2013, CPGS randomly surveyed urban dwellers in 30 Chinese large cities (all provincial capitals except Lhasa of Tibet). The effective sample size is 6,257.

and voice their various policy needs (Figure 2).

These empirical regularities echo scholars' earlier claims that the Chinese state sets up institutions that aim to solicit information from the citizens in an orderly and peaceful manner so as to sustain its rule (Nathan, 2003).[2] Botero, Ponce and Shleifer (2013) recently argue that, in the absence of competitive elections, the quality of authoritarian governance is tied to citizens' complaints. Along with this line of thought, this paper emphasizes the role of "deliberative institutions," institutions that are set up to allow ordinary citizens to speak, in sustaining government responsiveness and strengthening authoritarian rule.[3] Deliberation is regarded as not only a process of hierarchical communication both between the government and the citizens and among the citizens, but also as a way to aggregate citizens' preferences in a non-democratic setup.

This paper thus develops a game theoretic model that combines hierarchical communication (deliberation) with strategic government-policy responses. It attempts to answer one fundamental question: how an authoritarian government uses deliberative platforms to manipulate information in the public domain and respond to public opinion with policy changes, alleviating the pressure of citizens' collective action.

A basic problem faced by an authoritarian government is that it is uncertain about public opinion. For example, the government is not sure what fraction of the citizens are angry with the current policy. The citizens may not precisely know this either. If a majority of the citizens clearly know that they are all dissatisfied,

[2]Nathan codes the term "authoritarian resilience" to describe authoritarian states' strategies of adapting in new environments and seeking survival. He points out that "input institutions," such as letters-and-visits departments (*xinfang ju*), local people's congresses, and the mass media, play an important role in linking the Chinese authoritarian state and the society. Our definition of "deliberative institutions" is slightly narrower than his.

[3]Understanding patterns of government responsiveness has been one of the key issues in political economy (Besley and Burgess, 2002).

a credible threat of collective action can force the government to change the policy or even to step down. In order to prevent this from happening, the government has two options: (1) to improve the policy just enough to avoid the protest, and (2) to make the citizens believe that the collective action is not likely to be successful. Carefully designed deliberative institutions may help the government achieve these two tasks simultaneously in an efficient way.

To analyze the problem, we introduce a simple framework with one government actor and two citizens. In the baseline model, the government first decides whether to open deliberation and let the citizens speak. If permitted, each citizen sends a message at no cost (i.e., cheap-talk). These messages are public information.[4] Messages that correspond to a higher level of dissatisfaction as we label "complaints." The government then chooses a policy. After viewing the policy, the citizens simultaneously decide whether to participate in collective action demanding their desired policy.[5] The model is consistent with the empirical pattern reported by scholars that complaint-making to the Chinese state is highly individualized, issue-based and targets specific local-level government agencies or officials (e.g., Shi, 1997; Tsai and Xu, 2013).

We emphasize that deliberation is a process of hierarchical information transmission. It enables both vertical communication between the citizens and the government, and horizontal communication among the citizens.[6] Horizontal com-

[4]In Section 4, we relax the assumption of public deliberation by allowing the government to choose the degree of horizontal openness.

[5]The collective action can be either a small-scale protest, or a revolutionary social movement that leads to a regime change.

[6]This paper shares a similar insight as Farrell and Gibbons (1989), who investigate "cheap talk with two audiences." A key difference is that in this paper, the two receivers (the government and the other citizen) take action sequentially rather than simultaneously. Our model can also be understood as a veto bargaining game with pre-bargaining communication. The proposer is able to control the information inflow while the two citizens need to coordinate on exercising the "veto power."

munication, by making private information about individual preferences public, can either coordinate or disorganize citizens in collective action. We call these the *coordination* effect and the *disorganization* effect, respectively. As horizontal communication takes place, vertical communication, on the other hand, enables the government to respond to the fluctuating public opinion and reduce the risk of citizens' collective action by meeting their policy wishes. By allowing the government to condition policy on the vertical information flows, deliberation allows the government to preempt the protests that are caused by horizontal learning. Thus, the strategic response to vertical information flows mitigates the cost of horizontal learning. This sometimes tilts the cost and benefit of deliberation in favor of openness. It offers a direct justification for the wisdom in *Guo Yu* in our epigraph: a stifled people is like a blocked river; it is extremely dangerous if the ruler does not know that they are angry because they are not allowed to speak.

As a result, we show that the government's net gain from opening public deliberation can be decomposed into three effects. The first is the effect of vertical communication, which is positive, since the government can always make good use of the citizens' private information without its constraint being tightened. The second is the disorganization effect in horizontal communication, which is positive. The third one is the coordination effect in horizontal communication, which is negative.

The net value of the latter two effects can be either positive or negative, depending on whether the disorganization effect dominates the coordination effect. A citizen's incentive to join a protest depends on her perception of the other's participation. When citizens hold prior beliefs that it is very likely that the other will participate, and deliberation verifies it, the reinforcement of incentives to join the protest is miniscule. As a result, the coordination effect is small. If the deliberation,

however, falsifies the initial perception, citizens will get highly discouraged from protesting. Hence the disorganization effect is relatively larger. Similar reasoning applies to the case when the prior stands in favor of the government and citizens are perceived to be heterogeneous. In this case, the coordination effect dominates the disorganization effect. Thus, the net gain of horizontal communication is positive if and only if the initial perception does not stand in favor of the government.

In addition to beneficial vertical communication, our model suggests that horizontal communication may also do the government a favor by disorganizing the citizens. As channels of horizontal communication become more open, newly available information floating around in the public domain carries more weight than the initial perception in shaping citizens' beliefs. Hence, a citizen is more likely to be discouraged from participating in collective action when she finds that not many people file complaints. By emphasizing the disorganization effect of horizontal communication, we challenge the traditional idea that an authoritarian government unconditionally dislikes openness and the exchange of information among the citizens (Bueno de Mesquita and Downs, 2005; Hollyer, Rosendorff and Vreeland, 2011, 2013).

Based on the key tradeoff in the hierarchical communication, we also show that the relationship between regime openness and the threat of collective action (the government suffers from a successful protest) might be non-monotonic. When the threat is relatively small compared with the policy adjustment cost, the government mainly cares about its net payoff from horizontal communication (because it will never adjust policy anyway). In this case, an increase in the threat may lead to less freedom of expression since the regime is more concerned about coordination among the citizens as successful collective action becomes more damaging. However, when the threat is relatively large, more threat, in fact, forces the government to listen

to the citizens more carefully such that it improves the policy when needed. In this case, the policy adjustment effect from vertical communication dominates. This non-monotonic relationship challenges the conventional wisdom that political liberalization is almost always increasing in the threat of citizens' collective action (e.g. Acemoglu and Robinson 2000, 2001).

Due to the presence of the disorganization effect, we show that whenever the government prefers public deliberation to no communication, it also strictly prefers public deliberation to private polling. This result reinforces our basic intuition about the disorganization effect. By prohibiting citizens from communicating with each other, the government loses its chance to disorganize them. Another way to interpret the result is to think of public deliberation as the government's commitment not to manipulate information in the public domain. Therefore, messages received by citizens are persuasive and the resulting disorganization effect is large. Yet, with private polling, such an effect does not exist.

We also investigate two appealing institutional arrangements that strictly enhance the government's payoff. These two mechanisms further illustrate that certain forms of commitment help the government take advantage of citizens' learning and make the government better off. We refer to the first one as a *responsiveness* mechanism, in which the government shuts down direct horizontal communication (as if it conducts private polling) and commits to a level of responsiveness to the citizens' complaints. Such a mechanism reveals some information to the citizens through government policy changes. When the status quo is kept, based on the responsive rate, each citizen infers the other citizen's preference, and then decides whether to join a protest. A higher responsive rate makes a citizen less likely to believe that the other citizen also desires a policy change. Thus, the responsiveness mechanism plays the role of garbling information that is similar to the second

mechanism, which we refer to as a *censorship* mechanism. In such a mechanism, the government allows a certain form of horizontal communication but censors the information received from the citizens following a committed rule.[7] We show that the two mechanisms at their optimum levels are, in fact, equivalent since both of them feature garbling information in certain ways. Both mechanisms create strategic uncertainties among citizens by taking advantage of their horizontal learning, thus providing an opportunity for the government to reshape citizens' beliefs to its own advantage.

The welfare comparison between public deliberation and private polling, as well as the mechanism design exercises, shows that certain commitment devices can help the authoritarian government reap the benefits from the hierarchical communication while reducing its risks. In practice, there are many ways that the government can constrain itself from over-manipulating public opinion. Institutionalization of certain deliberative processes, by making credible commitments to the citizens, is one of them. The moderate responsiveness mechanism helps explain why the Chinese government sets up bureaucratic agencies to deal with citizens' complaints but only a small fraction of the petitioners achieve their goals. The partial but not full govermental responsiveness reflects the central government's intention to take advantage of citizens' learning. This suggests that even if the central government is able to perfectly overcome the agency problems with local bureaucracies, it lacks the incentive to hold them fully accountable. This is consistent with the empirical findings that the official appeal system helps mediate social conflicts but works rather "inefficiently" – only popular claims are addressed by the authorities (Cai, 2004; Chen, 2012).

[7]In such a mechanism, the government can be seen as an information "mediator" who truncates the pieces of information that it thinks are potentially dangerous (Meirowitz, Morelli, Ramsay and Squintani, 2013).

The effect of commitment in using horizontal learning, including the effect of censorship mechanism, on the other hand, explains why the Chinese government allows private companies to set up online platforms where people can speak freely as long as strict censorship protocols are followed. The model makes predictions that are consistent with the following empirical findings: (1) the government is not afraid of being criticized online but is extremely sensitive to posts that may rally collective action (King, Pan and Roberts, 2013), and (2) whenever an issue grabs headlines on the Internet and spurs popular anger among the citizens, the state reacts very quickly to appease the anger by disciplining exposed local officials accused of misbehaving or by improving policies (Distelhorst, 2013).

This paper contributes to the recent literature on the mechanics of authoritarian rule. For example, Gehlbach and Keefer (2011) and Boix and Svolik (2013) argue that institutionalization of the ruling party creates credible commitment for investors, and helps solve the power-sharing problem within the elite circle, respectively. Little (2012) and Egorov and Sonin (2012) suggest that the state uses controlled elections to signal its strength and deters collective action. Lohmann (1993) and Lorentzen (2013*a*) illustrate how the state can learn from citizens' extreme political acts, including riots. Egorov, Guriev and Sonin (2009) and Lorentzen (2013*b*) explain how authoritarian regimes use semi-independent media to improve the quality of governance in moral hazard environments. This paper, however, focuses on citizens' non-electoral political participation as a contributing factor of government responsiveness. Different from studies in which the government is assumed to care about citizens' well-being directly, our model provides a microfoundation that explains the government's incentive to adjust policies.

This paper is also closely related to work that investigates the government control of information and citizens' collective action (e.g., Acharya 2012; Little 2012,

2013; Egorov and Sonin 2012; Edmond 2013; Gehlbach and Sonin 2013; Lorentzen 2013*b*). For example, Gehlbach and Sonin (2013) show that government control of the media in democratic countries helps mobilize citizens for political objectives, and Lorentzen (2013*b*) argues that an authoritarian government allows a certain degree of media freedom to deter corruption at the cost of social instability. Our paper is different from most of these works in two aspects. First, in our model, the citizens learn each other's preference instead of an unknown common fundamental, thus deliberation serves as a preference aggregation mechanism. Second, as in Little (2012, 2013), we assume that the government does not know the state of the world that affects citizens' incentives to participate in collective action. This difference is significant. In models where the government knows the fundamental, the government's choice is whether to reveal the information truthfully under different circumstances. In our model, however, the choice is between a pooling equilibrium, where the government learns nothing, and an informative one, where the government learns a lot. The first case is naturally applicable to issues such as government transparency and media supervision, while our model more directly targets questions about how government behavior tracks public opinion.

This paper thus makes three contributions to the literature. It is among the first to directly investigate both vertical and horizontal communication in a unified, simple framework. Second, it demonstrates the disorganization effect of communication which strengthens authoritarian governance. Third, it shows how deliberation aggregates citizens' preference aggregation through non-electoral political participation.

The arrangement of the paper is as follows. Section 2 introduces the basic framework. Section 3 investigates the main comparative statics. Section 4 compares public deliberation with private polling, and proposes two institutional arrangements

that provide the government with a higher payoff than public deliberation, private polling and no communication. Section 5 concludes and discusses the implications of the model.

2 A Benchmark Model

In this section, we introduce the basic framework and explain the mechanics of vertical and horizontal communication with a simple benchmark model. The Appendix presents the proofs. An online Supplementary Appendix explores several discussions and extensions of the benchmark model.[8]

2.1 Setup

Players and policy preferences. There are three players: a government, citizen 1 and citizen 2. The government has two policy options to choose from, $x \in \{Q, R\}$. We call Q the status quo policy and R the reform policy. It costs the government $\mu > 0$ to implement the reform policy relative to the status quo (hereafter μ is called the policy adjustment cost), as if the government draws a positive payoff gain μ from Q relative to R. We denote $Q = 0$ and $R = 1$ for notational convenience.

The two citizens can be one of two types, a *non-activist*, who is indifferent between reform and the status quo, or an *activist*, who strictly prefers reform to the status quo. Citizen i's type is denoted as $\omega_i \in \{\underline{\omega}, \overline{\omega}\}$, with $\underline{\omega}$ and $\overline{\omega}$ representing non-activist and activist, respectively. We normalize a citizen's policy gain from

[8] In the Appendix, we provide the proofs for most of the results based on a more generalized model, of which the benchmark model is a special case. The key difference between the two models is that in the benchmark model, citizens' collective-action gain is purely from policy outcomes, while in the more generalized model, citizens' collective-action gain takes a flexible form, thus allowing richer interpretations. For example, citizens' collective action can be motivated by psychological factors such as grievance (Passarelli and Tabellini, 2013).

the status quo to zero, no matter what type he or she is, i.e.,

$$u_i(Q;\omega_i) = 0, \quad i = 1, 2. \tag{1}$$

A non-activist is indifferent between the reform policy and the status quo:[9]

$$u_i(R;\omega_i = \underline{\omega}) = \underline{L} = 0, \quad i = 1, 2. \tag{2}$$

An activist gets strictly positive payoff from the reform policy:

$$u_i(R;\omega_i = \overline{\omega}) = L > 0, \quad i = 1, 2, \tag{3}$$

where L is common knowledge. Figure 3 shows the policy preferences of the government and the two types of citizens.

[Insert Figure 3]

Since the public opinion fluctuates over time, the citizens' types ω_1 and ω_2 are unknown to the government. However, both the government and the citizens share a common prior that a citizen is an activist with probability one-half:[10]

$$\Pr(\omega_i = \overline{\omega}) = p = \frac{1}{2}, \quad i = 1, 2. \tag{4}$$

The two citizens' types are (potentially) correlated. If one of the citizens is an

[9]For simplicity, we assume $\underline{L} = 0$ for simplicity. When $\underline{L} < 0$, the main results of the paper still hold. When $0 < \underline{L} < \overline{L}$, an informative equilibrium can be sustained if and only if the cost of complaining is appropriately designed. Otherwise, no information may be induced in deliberation since the low type always pretends to be the high type.

[10]For simplicitye, we assume $p = \frac{1}{2}$. The main results do not qualitatively depend on p. Allowing different players to have different priors, as long as they are public information, does not change the results either.

activist, with probability γ, the other one is also an activist, i.e.,

$$\omega_j|_{\omega_i=\overline{\omega}} = \begin{cases} \overline{\omega} & \text{with probability} \quad \gamma \\ \underline{\omega} & \text{with probability} \quad 1-\gamma \end{cases}, \quad i,j \in \{1,2\},\ i \neq j. \tag{5}$$

γ represents citizens' preference correlation. Both the prior $p = \frac{1}{2}$ and preference correlation γ are common knowledge and capture the public information in the society.[11] Lower preference correlation γ implies that *ex ante* citizens are more likely to be heterogeneous.

Timing and actions. The timing of actions is as follows.

Period (0) *Institutional design.* The government chooses whether or not to open public deliberative platforms to let citizens speak, $\alpha \in \{0,1\}$. When $\alpha = 0$, a citizen's complaint will not be heard by the government or her fellow citizen. On the contrary, when $\alpha = 1$, citizens are allowed to send messages and their messages will be heard by both the government and the other citizen.[12]

Period (1) *Public deliberation.* If allowed (i.e., $\alpha = 1$), each citizen sends a message $m_i \in \{0,1\}$ to the government at no cost. The message is publicly observable. We interpret $m_i = 1$ as complaining and $m_i = 0$ as abstaining. If $\alpha = 0$, this period is skipped.

Period (2) *Policy adjustment.* The government chooses a policy, $x \in \{Q, R\}$, which is publicly observable.

Period (3) *Collective action.* Each citizen simultaneously decides whether to participate in a popular protest ($a_i = 1$) or not ($a_i = 0$).

The time-line of the benchmark model is also summarized in Figure 4.

[11] The distribution of $\omega_j|_{\omega_i=\underline{\omega}}$ is characterized in equation (A4) in the Appendix.

[12] In Section 4, we allow the government to choose the extent to which a citizen's complaint is heard by one another.

[Insert Figure 4]

Collective action and payoffs. Each player's payoff consists of two parts: a payoff at the policy adjustment stage, and a payoff at the collective-action stage. We denote them as the "policy payoff" and the "collective-action payoff," respectively, although the latter can also be seen as policy driven. For simplicity, we assign equal weight to the two payoffs.[13]

Successful collective action requires both citizens' participation. In this case, the reform policy gets implemented and the government suffers $\rho_2 > 0$.[14] If only one of the citizens participates, the government suffers a cost $\rho_1 > 0$, which is smaller than ρ_2. Moreover, with probability λ, the individual protest is successful and the reform policy is implemented; with probability $(1 - \lambda)$, it is not successful and the original policy x remains unchanged. If neither citizen participates, no policy change happens and the government suffers no cost. Citizen i's collective-action payoff is represented by:

	participate (j)	abstain (j)
participate (i)	$u_i(R;\omega_i) - k_i$	$\lambda u_i(R;\omega_i) + (1-\lambda)u_i(x;\omega_i) - k_i$
abstain (i)	$\lambda u_i(R;\omega_i) + (1-\lambda)u_i(x;\omega_i)$	$u_i(x;\omega_i)$

where k_i is citizen i's cost of participating in collective action. k_i is i's private information and is only known to her after she observes the government's policy x.[15] k_i is assumed to be independent and identically distributed between 0 and

[13]The results are qualitatively the same if we assign different weights. In the extreme case, if we put zero weight on the policy payoff, the model is reduced to a classical veto bargaining game.

[14]Depending on the magnitude of ρ_2, we can interpret collective action in different ways. It can be a small-pscale protest demanding the government to change a particular policy or to punish a misbehaved local official, as often happens in China. It can also be a social movement aiming at a regime change, after which citizens or a new government implement the reform policy.

[15]In fact, we can show that, even if citizens know their private costs of collective action in the deliberation stage, the outcome induced in any symmetric cut-point equilibrium is the same as that when they do not know the costs. The intuition is that a citizen with a high collective-action

1 with cumulative distribution function $F(\cdot)$. We assume $F(\cdot)$ is weakly concave; and $f(k) = F'(k) > 0, \forall k \in (0,1)$.[16]

The government's total payoff also consists of two parts: a policy implementation cost and a cost from collective action. We summarize its total payoff as follows (recall $Q = 0$ and $R = 1$):

	participate (j)	abstain (j)
participate (i)	$-x\mu - (1-x)\rho_2$	$-x\mu - (1-x)\rho_1$
abstain (i)	$-x\mu - (1-x)\rho_1$	$-x\mu$

We further make two assumptions making a successful protest attractive.[17] First, an activist's gain from the reform policy is larger than the upper bound of the collective-action cost, i.e, $L > 1$. Second, the probability that an individual challenge succeeds is relatively small, i.e.,

$$\lambda < \min\{\frac{1}{L}, 1 - \frac{1}{L}\}. \tag{6}$$

Denote $A \equiv (1-\lambda)L$, which is the payoff gain of joining a protest (excluding the protest cost) provided that the other citizen also participates. Similarly, $B \equiv \lambda L$, is the payoff gain when the other citizen does not participate. Hence, we have $0 < B < 1 < A$.[18]

cost always has an incentive to pretend to be of low cost so as to persuade the other citizen to join the protest. As a result, in any symmetric cut-point equilibrium, cheap-talk produces no information of the collective-action cost that would change the equilibrium outcome.

[16]It can be verified that the uniform distribution and any distribution with a cumulative distribution function $F(k) = k^\delta$ $(0 < \delta < 1)$ satisfy this property. The concavity of the distribution is used merely to guarantee the unique prediction in the collective action. Without this assumption, we may need to deal with the problem of multiple equilibria, although the properties in the equilibrium we focus on are still valid.

[17]The two assumptions are used merely to simplify the presentations of the theoretical argument. The main results remain unchanged when the two assumptions are relaxed.

[18]We provide technical interpretations for these conditions in the Supplementary Appendix.

The equilibrium notion is (Perfect Bayesian) *Nash Equilibrium*. Since there may exist multiple equilibria as in other cheap-talk/signaling games, we focus on the equilibrium in which citizens truthfully reveal their types (preferences) when allowed. We first pin down the equilibrium features at the collective-action stage based on the conjecture that such an equilibrium exists. Then, we use those properties to verify its existence by checking incentive compatibility constraints of citizens at the deliberation stage. Third, we investigate the government's optimal choice of whether to allow public deliberation.

Without loss of generality, we assume that when the government is indifferent between allowing and forbidding deliberation, it chooses the latter ($\alpha^* = 0$). In addition, when the government is indifferent between implementing the reform policy and keeping the status quo, it keeps the status quo ($x^* = Q$).[19]

2.2 Equilibrium Characterization

The collective-action stage. The following lemma (as well as a more technical version Lemma 2 in the Appendix) shows that in any truth-telling equilibrium, the citizens' strategy at the collective-action stage is uniquely determined.[20] Specifically, when reform policy is implemented, no one protests; when the status quo policy is implemented, a non-activist never protests, whereas an activist protests

[19]Lemma 3 in the Appendix suggests that the fully revealing equilibrium still exists without these two assumptions. Thus, they do not affect the incentives for deliberation, and are used just to exclude discussions of additional equilibria that arise with probability zero within the parameter space.

[20]Without the standard *common-value global games* setup, we still get the uniqueness feature at the collective-action stage. As implied by Morris and Shin (2006), whether the prediction for collective action in both common value and private value games is unique depends on how we technically parameterize the payoffs and the distribution of uncertainty. The technical approach we use, i.e., by incorporating private costs, is similar to Palfrey and Rosenthal (1985), that Morris and Shin (2006) call a *private-value interaction/global game*. The technical details of Lemma 1 are shown in Lemma 2 in the Appendix.

if and only if the realized cost of collective action is relatively small.

Since neither the government nor the other citizen observes a citizen's cost, citizens' equilibrium behaviors of whether to join a protest appear random both to the government and to each other. In the following, we only present the probability by which a citizen joins collective action from the *ex ante* point of view, while leaving the technical characterization of the cut-point strategy in Lemma 2. When the citizens are not allowed to speak, the probability of an activist participating in collective action is increasing in γ. This is because an activist understands that with bigger γ, the likelihood that her fellow citizen also demands reform is higher, even though the two cannot communicate with each other directly. In the extreme case when the two citizens have purely opposite preferences, i.e., $\gamma = 0$, the probability of an activist joining a protest reaches the minimum. As γ goes up, it increases until reaching the highest value 1.

To state the following lemma appropriately, we define a function $p_0(\cdot)$ in Equation (A8) in the Appendix. $p_0(\gamma)$ is the probability that an activist protests when deliberation is not allowed.

Lemma 1 *(Characterizing the collective-action stage) In any equilibrium in which both citizens truthfully reveal their types when they are allowed to speak, each citizen uses a cut-point strategy characterized in Lemma 2 in the Appendix. Specifically, (1) an non-activist never protests; (2) under the reform policy, an activist never protests; (3) under the status quo policy, if deliberation is not allowed, the probability that an activist protests equals $p_0(\gamma)$; and (4) under the status quo policy, if deliberation is allowed, the probability that an activist protests equals 1 or $p_0(0)$ when she faces an activist or a non-activist, respectively. $p_0(\gamma)$ is strictly increasing in γ when $\gamma \leq \gamma_0 = \frac{1-B}{A-B}$; $p_0(\gamma) = 1$, when $\gamma \geq \gamma_0$.*

When an activist sees that the other citizen claims that he is also an activist by filing a complaint, she joins a protest for sure since she knows the other one will do the same thing, and the collective action will be successful. If she sees the other citizen abstain, she infers that he is a non-activist and chooses to join a protest with probability $p_0(0)$. However, if deliberation is not allowed, an activist can only condition her behavior on public information. The probability of her joining a protest $p_0(\gamma)$ increases with γ since she knows with bigger γ the prospect that the other citizen is also an activist is bigger, and the collective action is more likely to be successful. Therefore, a bigger γ implies more credibility of citizens' collection action. By informing the government of the citizens' private information, deliberation helps the government to identify the actual threat from collective action. We summarize an activist's decision rule to join a protest without deliberation in Figure 5.

[Insert Figure 5]

Lemma 1 also helps us to understand the role of horizontal communication in facilitating or impeding collective action. Compared with no communication, with deliberation, an activist increases her probability of joining a protest from $p_0(\gamma)$ to 1, when she finds the other one is also an activist ($m_{-i} = \overline{\omega}$) and decreases her probability of protest from $p_0(\gamma)$ to $p_0(0)$ when she finds the other citizen is a non-activist ($m_{-i} = \underline{\omega}$). We illustrate this effect in Figure 6.

[Insert Figure 6]

In Figure 6, for $\gamma = \gamma_L$ or $\gamma = \gamma_H$, the non-dotted arrow represents an activist's change of behavior at the collective action stage (from no communication to public deliberation) when she finds the other citizen is an activist. This suggests that

the citizens (two activists) are better coordinated through deliberation and the government is worse off. Similarly, the dotted arrow represents the change when she finds the other citizen is a non-activist. This means the activist is discouraged from joining a protest when she knows the other one will not join, and in this case, the government is better off.

The preference correlation γ can be deemed as a piece of public information that shapes both the citizens' initial beliefs of each other's preference, and the government's initial belief about the citizens. As γ becomes smaller (say, from γ_H to γ_L), it is more favorable to the government. As the citizens become more heterogeneous, in the absence of communication, they are less likely to coordinate with each other under the status quo policy. When deliberation is allowed, a citizen's initial belief (based on γ) of the probability that the other is an activist is updated either downward to zero, when she sees no complaints, or upward to one, when she sees the other's complaint. Therefore, Figure 6 also suggests that a bigger γ means that the initial belief is less likely to be upward when it is updated by deliberation. This implies larger discouragement for collective action when an activist faces a non-activist, and a smaller encouragement when she faces an activist when deliberation is allowed.

The government's decision to open deliberation. Since no one will protest against the reform policy, in front of the decision of opening deliberation, the government simply compares the policy adjustment cost μ and the expected cost from the citizens' collective action.

Suppose $G_1(x, \omega_1, \omega_2)$ is the government's payoff in every possible situation (ω_1, ω_2) when it allows deliberation and it chooses a policy $x \in \{Q, R\}$; $G_0(x, \omega_1, \omega_2)$ is its payoff in every possible situation (ω_1, ω_2) when it forbids deliberation and

chooses a policy $x \in \{Q, R\}$.

We first discuss the case when deliberation is allowed. Remember, if the government chooses the reform policy, no one protests. If both citizens are activists, i.e., $\omega_1 = \omega_2 = \overline{\omega}$, they know each other's type through deliberation and protest against the status quo policy with probability 1. The government's payoff therefore is:

$$G_1(x, \omega_1, \omega_2) = -(1 - x)\rho_2 - \mu x. \tag{7}$$

If both citizens are non-activists, i.e., $\omega_1 = \omega_2 = \underline{\omega}$, they do not protest against the status quo policy. The government's payoff is:

$$G_1(x, \omega_1, \omega_2) = -\mu x. \tag{8}$$

If there is one activist and one non-activist, i.e., $\omega_1 \neq \omega_2$, the activist protests against the status quo policy with probability $p_0(0)$. Hence, the government's payoff is:

$$G_1(x, \omega_1, \omega_2) = -(1 - x)\rho_1 p_0(0) - \mu x. \tag{9}$$

Similarly we can write down the government's payoff when deliberation is not allowed. The government's payoffs in each of the three cases under the two circumstances ($\alpha = 0$ and $\alpha = 1$) and their differences are summarized in the following table.[21] For the moment, we use the same notation x for the policy adjustments when the deliberation is both allowed and not allowed, thus implicitly assuming the government policies under the two circumstances are the same.

[21]The equivalent expressions of the government payoffs are Equations (A1) and (A2) in the Appendix.

	$\omega_1 = \omega_2 = \underline{\omega}$	$\omega_1 = \omega_2 = \overline{\omega}$	$\omega_1 \neq \omega_2$
$\alpha = 1$	$-\mu x$	$-(1-x)\rho_2 - \mu x$	$-(1-x)\rho_1 p_0(0) - \mu x$
$\alpha = 0$	$-\mu x$	$-(1-x)\widehat{\rho} - \mu x$	$-(1-x)\rho_1 p_0(\gamma) - \mu x$
$G_1 - G_0$	0	$-(1-x)(\rho_2 - \widehat{\rho}) < 0$	$(1-x)\rho_1[p_0(\gamma) - p_0(0)] > 0$

where $\widehat{\rho} = [\rho_2 p_0(\gamma)^2 + \rho_1 2p_0(\gamma)(1-p_0(\gamma))] < \rho_2$.

We see that, without considering the learning effect of its own, the government benefits from deliberation when the two citizens are of different types and loses from deliberation when both citizens are activists. In other words, public deliberation might be beneficial to government even if it does not adjust policies according to what it learns from the citizens.

Horizontal communication has two possible effects on the government's net payoff from opening deliberation: the *coordination* effect, when the citizens find out that they both are activists, i.e., $\omega_1 = \omega_2 = \overline{\omega}$, and the *disorganization* effect, when the citizens realize that they are of different types, i.e., $\omega_1 \neq \omega_2$.

Suppose x^* is the government's optimal policy choice when deliberation is not allowed, i.e., $x^* \in \arg\max_{x \in \{Q,R\}} E[G_0(x, \omega_1, \omega_2)]$. The coordination effect is formally defined as the government's net gain resulting from citizens' horizontal learning when they are of the same type, assuming the government sticks to x^*, i.e.,

$$\sum_{w_1 = w_2} \Pr(\omega_1, \omega_2)[G_1(x^*, \omega_1, \omega_2) - G_0(x^*, \omega_1, \omega_2)], \tag{10}$$

which equals $-\frac{1}{2}(1-x^*)\gamma(1-p_0(\gamma))[(1+p_0(\gamma))\rho_2 - 2p_0(\gamma)\rho_1]$ after simplification and is always non-positive.

Similarly, the disorganization effect is formally defined as a net gain from horizontal learning when citizens are different types, assuming the government sticks to x^*, i.e.,

$$\sum_{w_1 \neq w_2} \Pr(\omega_1, \omega_2)[G_1(x^*, \omega_1, \omega_2) - G_0(x^*, \omega_1, \omega_2)], \tag{11}$$

which equals $(1-\gamma)(1-x^*)(p_0(\gamma)-p_0(0))\rho_1$ after simplification and is always non-negative.

Vertical communication affects the government's payoff through a direct learning effect. The government may also gain from opening public deliberation since it learns from citizens' and adjusts policy when it finds that the citizens pose an *actual* threat to its rule. The *policy adjustment* effect from vertical communication is formally defined as the government's net gain from learning the citizens' preferences through deliberation, assuming that the citizens already know each other's preferences, i.e.,

$$E[\max_x G_1(x,\omega_1,w_2) - G_1(x^*,\omega_1,\omega_2)], \tag{12}$$

which is always non-negative since there is no loss to obtain additional information that does not change the constraints the government faces.

Hence we get the following **Hierarchical Communication Identity**, that shows the government's payoff difference between allowing and forbidding deliberation can be decomposed into the above three effects, i.e.,

$$\begin{aligned}
&E[\max_x G_1(x,\omega_1,\omega_2)] - G_0(x^*,\omega_1,\omega_2)] \\
=\ &\underbrace{E[\max_x G_1(x,\omega_1,w_2) - G_1(x^*,\omega_1,\omega_2)]}_{\text{the policy adjustment effect from vertical communication}} \\
+\ &\underbrace{\sum_{w_1=w_2} \Pr(\omega_1,\omega_2)[G_1(x^*,\omega_1,\omega_2) - G_0(x^*,\omega_1,\omega_2)]}_{\text{the coordination effect from horizontal learning}} \\
+\ &\underbrace{\sum_{w_1\neq w_2} \Pr(\omega_1,\omega_2)[G_1(x^*,\omega_1,\omega_2) - G_0(x^*,\omega_1,\omega_2)]}_{\text{the disorganization effect from horizontal learning}}
\end{aligned} \tag{13}$$

We summarize the above results in the following proposition.

Proposition 1 ***The Hierarchical Communication Identity*** *(13) always holds. Namely, the government's net gain of opening deliberation can be decomposed into three effects: (1) the policy adjustment effect (a direct informational gain from vertical communication, provided that the citizens already know each other's type) that is non-negative; (2) the coordination effect (a net gain from horizontal communication when the citizens are of the same type, provided that the government sticks to the policy choice without deliberation) that is non-positive; and (3) the disorganization effect (a net gain from horizontal communication when citizens are of different types, provided that the government sticks to the policy choice without deliberation) that is non-negative.*

Having clarified the basic tradeoff of the government's decision, we directly write down the difference in the government's payoff between allowing and forbidding deliberation. When deliberation is not allowed, the government gets $\max\{-M, -\mu\}$, where

$$M = \frac{1}{2}\gamma p_0(\gamma)^2\rho_2 + p_0(\gamma)\rho_1(1 - \gamma p_0(\gamma)) \tag{14}$$

is the government's expected loss from the citizens' collective action. The "max" function represents the government's choice whether to implement the reform policy. When deliberation is allowed, the government will get $\max\{-\rho_2, -\mu\}$ if it sees two complaints (of both activists), get $\max\{-p_0(0)\rho_1, -\mu\}$ if it sees only one complaint and get 0 if it receives no complaints. The difference in the government's payoff therefore is:

$$\min\{M, \mu\} - \frac{1}{2}\gamma \min\{\rho_2, \mu\} - (1 - \gamma)\min\{p_0(0)\rho_1, \mu\}, \tag{15}$$

which is dependent on γ, μ, ρ_1 and ρ_2. The following existence result not only

characterizes the government's optimal decision of opening deliberation, but also confirms our initial conjecture that truthfully revealing their types is incentive compatible for the citizens.

Proposition 2 ***(Existence of a truth-telling equilibrium)*** *(1)A truth-telling equilibrium exists, in which whenever $\alpha = 1$, all citizens truthfully reveal their preferences. (2) In this equilibrium, the government allows deliberation ($\alpha^* = 1$) if and only if $\min\{M, \mu\} - \frac{1}{2}\gamma \min\{\rho_2, \mu\} - (1-\gamma)\min\{p_0(0)\rho_1, \mu\} > 0$, where $M = \frac{1}{2}\gamma p_0(\gamma)^2\rho_2 + p_0(\gamma)\rho_1(1 - \gamma p_0(\gamma))$.*

(We provide a proof for a more technical version of it, Lemma 3, in the Appendix.)

3 Comparative Statics

In this section we explore the main comparative statics of the benchmark model.

3.1 Preference correlation

We have established the existence of a fully revealing equilibrium (Proposition 2) and have shown that the prediction for the collective-action stage is unique (Lemma 1). Based on these results, we are now able to investigate the effect of citizens' preference correlation on openness.

We can think of the preference correlation parameter γ as the credibility of the threat of collective action when deliberation is not allowed. The policy adjustment effect from vertical communication increases monotonically with γ, since as γ increases, with greater chances deliberation can help the government identify potential collective action risks and adjust policies when these actions are needed.

We are interested in γ since it is a determinant of the citizens' initial belief of each other's preference. Whether the disorganization effect dominates the coordination effect in horizontal communication depends on the citizens' initial belief. If γ is small, an activist's prior that her fellow citizen is also an activist is small. When deliberation shows that he is indeed an activist, her incentive to join a protest increases dramatically; otherwise her incentive to join a protest does not decrease too much. In this case, the coordination effect dominates. Similar logic applies to the case when γ is large, and the disorganization effect dominates.

The traditional idea is that an authoritarian government does not want the citizens to communicate with each other, since communication makes them better coordinated (Hollyer, Rosendorff and Vreeland, 2011, 2013). Consistent with this idea, in our model, the coordination effect discourages the government from opening deliberation when γ is relatively small. Although the benefits from the disorganization effect and the policy adjustment effect from vertical communication still exist, the coordination effect dominates as the prospect of a successful protest is slim in the absence of horizontal communication. As a result, the government forbids deliberation. Under this circumstance, policy efforts made by the government are also limited since there is no great pressure from the citizens.

However, when γ is large, the coordination effect is dominated by the other two effects. Vertical communication becomes more valuable to the government as γ increases since the citizens' collective action becomes more credible when they are not allowed to speak, thus imposing a larger threat to the government. By allowing deliberation and listening to their messages carefully, the government can address the citizens' policy demand in a case-by-case manner and avoid potential collective action. Moreover, when γ increases, the horizontal communication may give the government a positive net payoff. In other words, openness creates a

greater opportunity to disorganize the citizens by informing them of one another's true type. Overall, the possibility of a protest is reduced. We summarize the argument in the following proposition.[22]

Proposition 3 ***(Citizens' preference correlation pushes openness)*** *Provided that* $\frac{F^{-1}(y)-B}{y}$ *is concave,*[23] *when* $\rho_2 \geq 2\rho_1$, $\rho_2 > c$, *and* λ *is sufficiently small, (in the truth telling equilibrium) the government's payoff function is U-shaped w.r.t. the citizens' preference correlation* γ, *and* $\exists \gamma^* \in (0, \gamma_0)$, *where* $\gamma_0 = \frac{1-B}{A-B}$, *such that*

$$\alpha^* = \begin{cases} 1 & \text{if } \gamma > \gamma^* \\ 0 & \text{if } \gamma \leq \gamma^* \end{cases} . \tag{16}$$

Figure 7 shows how γ affects the government's incentive to open deliberation through both vertical and horizontal communication.[24] It also suggests that even without the policy adjustment effect, it might still be optimal for the government to open deliberation when γ is relatively large, in which case the disorganization effect dominates the coordination effect.

3.2 The threat of collective action

We define the threat of collective action as the cost in which the government suffers from a successful protest ρ_2. Different from the popular view (Acemoglu and

[22] Although we only discuss the effect of γ given the prior $p = \frac{1}{2}$, in a slightly different case when the two citizens' types are independent and identical, we can also show that the prior p has qualitatively the same impact as γ on the government's incentive to open deliberation.

[23] The uniform distribution naturally satisfies this condition.

[24] The U-shape relationship is mainly driven by the coordination effect when γ is relatively small. The government's incentive to open deliberation first decreases then increases with γ due to two competing forces: the odds that the two citizens share the same policy demand, i.e., $Pr(\omega_1 = \omega_2 = \overline{\omega})$, which is increasing in γ, and their changes of behavior when deliberation is allowed, $1 - p_0(\gamma)$, which is decreasing in γ. It turns out that the latter dominates at least in terms of behavior.

Robinson, 2000, 2001), we show that the relationship between regime openness (in terms of allowing people to speak), and the threat of collective action is non-monotonic. The authoritarian state is willing to open deliberation when the threat is very small or relatively large compared with the policy adjustment cost.

To illustrate, let's first consider the case when the threat of collective action is relatively small, i.e., $\rho_2 < \mu$. Since policy adjustment is expensive to the regime, it never changes the status quo policy. As a result, vertical communication does not affect the government's payoff, and the government's decision of opening deliberation rests wholly on horizontal communication. The coordination effect decreases with ρ_2, as a successful protest becomes more damaging to the government, while the disorganization effect remains constant and positive. Combining these two together, it is clear that as the threat of collective action becomes smaller, the government gains from citizens' horizontal learning.

However, when the threat of collective action is relatively large, i.e., $\rho_2 > \mu$, the government has an incentive to open deliberation and implement the reform policy once informed that both citizens are activists. Under this circumstance, the policy adjustment effect dominates. The larger the threat is, the more the government gains from learning the citizens' preferences through vertical communication and adjusting the policy accordingly. We summarize this line of reasoning in the following proposition.

Proposition 4 ***(Non-monotonic relationship between the threat of collective action and openness)*** *Provided conditions in Proposition 10 in the Appendix, there exist two thresholds $\overline{\rho}, \underline{\rho}$ (with $\overline{\rho} > \mu > \underline{\rho} > \rho_1$) such that the government chooses to open deliberation if and only if the threat of collective action ρ_2 is very small ($\rho_2 < \underline{\rho}$) or large ($\rho_2 > \overline{\rho}$); the government prohibits deliberation when the threat is moderate, i.e., $\underline{\rho} \leq \rho_2 \leq \overline{\rho}$.*

In Figure 8, we show how the three effects, as well as the aggregate incentive of the government to open deliberation, change with the threat of collective action.

[Insert Figure 8]

The non-monotonicity originates from the possibility that $\rho_2 < \mu$. Recall that ρ_2 is the cost the government pays when both citizens join a protest, in which case, the reform policy is implemented. There are several explanations why ρ_2 could be smaller than μ. First, as in the more generalized model we analyze in the Appendix, citizens' collective action can be driven by factors other than their policy preferences, in which case there is no natural linkage between the policy adjustment cost μ and the government's suffering from collective action ρ_2. Second, if we assume the citizens' collective action is purely policy driven, it is possible that under the citizens' direct monitoring after a successful protest, the policy implementation cost is lowered.[25] Third, it is possible that when the government initially wants to implement the reform policy, it faces considerable bureaucratic backlash, which can be overcome when the public opinion is seen to be leaning toward the reform policy.

3.3 Other Comparative Statics

Policy rigidity, measured by the policy adjustment cost μ, captures the conflict of interest between the authoritarian government and the citizens. As the preferences of the government and the citizens diverge, naturally it would be harder for the citizens to convince the government to adopt their preferred policy, and the gov-

[25] One can think of a case in which the reform policy R can be successfully implemented only with probability π when the government initially wants to implement it. The threat of collective-action ρ_2 therefore could be smaller than μ/π.

ernment is less incentivized to pay attention to the citizens' voice. Following this logic, Proposition 5 shows that higher policy adjustment cost implies less openness.

Proposition 5 ***(Effect of policy rigidity)*** *When the threat of collective action ρ_2 is sufficiently large, γ and λ are sufficiently small, $\exists \mu^* \in (0, \rho_2)$ such that*

$$\alpha^* = \begin{cases} 1 & \text{if } \mu < \mu^* \\ 0 & \text{if } \mu \geq \mu^* \end{cases} . \tag{17}$$

(See Supplementary Appendix for the proof.)

The model also allows us to clarify the relationship between openness and policy responsiveness, both of which are strategic choices made by the government. We define openness (as opposed to closeness) as the government's choice of allowing public deliberation (as opposed to forbidding it), and define policy responsiveness as the probability that the government implements the reform policy at the policy adjustment stage. We show that openness does not have a monotonic causal impact on government responsiveness, although they are positively correlated. We provide the detailed discussion and the proof in the Supplementary Appendix and directly present the following proposition.

Proposition 6 ***(Openness and policy responsiveness)*** *Assuming that $\rho_2 > \mu > \rho_1$, in equilibrium we have: (1) policy responsiveness is positively correlated with openness; and (2) more openness does not necessarily lead to more responsiveness.*

In the Supplementary Appendix, we also extend the model to incorporate the fact that citizens may have different preferences over the directions or forms of reform, so that the government may take advantage of this to keep the status quo.

4 Designing Deliberative Institutions

In Section 2, we have already shown that the government's decision to open deliberations hinges on three competing effects: the policy adjustment effect, the disorganization effect and the coordination effect. The first two effects encourage the government to open deliberation while the third one discourages it. In this section, we slightly deviate from the standard signaling/cheap-talk setting and think of the government as an institutional designer. The basic idea is to show that certain forms of credible commitment enable the government to better utilize citizens' learning.

To illustrate the basic idea, we first compare the government's welfare under public deliberation and its welfare under private polling, in which citizens' messages can only be directly seen by the government. Under public deliberation, the government implicitly commits not to manipulate citizens' beliefs about each other, since when observing public opinion, the citizens also observe each other's preference. However, with private polling, such commitment disappears because the only channel by which citizens can "observe" each other's preference is the potentially informative policy adjustment of the government. In this case, without commitment, the government has an incentive to deviate from the fully separating equilibrium, which is equivalent to the equilibrium induced in public deliberation. We will show that the fully separating equilibrium (if it exists in private polling), by allowing citizens' horizontal learning, fully uses the disorganization effect, and

thus makes the government better off than any other outcome in the private poll. The welfare comparison between private polling and public deliberation shows that the government can do better with certain credible commitments not to manipulate citizens' learning.

We then proceed to propose two seemingly different mechanisms that can raise the government's payoff compared with the payoff under public deliberation. By committing itself to either of the two mechanisms, the government can reap the benefits of hierarchical communication while reducing its harm to the greatest possible extent. In both mechanisms, direct communication among citizens is not allowed. In the *responsiveness* mechanism, the government commits to a certain response rate to the citizens' complaints. The government implements the reform policy with a given probability only when both citizens file complaints. In the *censorship* mechanism, the government receives the citizens' messages, filters them, and selectively reveals the information to the public. The probability that the information is censored when both citizens file complaints is decided in advance. Both mechanisms share a feature of selective information disclosure that affects horizontal communication. In this way, the government can manipulate the citizens' beliefs of each other's preference, thus changing the probability of their participating in collective action. Last but not least, we show that the two mechanisms are essentially equivalent. At their respective optima, they induce the same equilibrium outcome. The government is strictly better off under the two mechanisms than with public deliberation, purely private polling with no commitment, and no communication.

The welfare comparison between the best mechanisms and private polling confirms the basic idea that the government can raise its payoff by constraining itself from manipulating public opinion. The welfare comparison between the best

mechanisms and the public deliberation implies that the government should not constrain itself too much. Although such a proposition slightly deviates from the implication of the first comparison, it does confirm the main theme of the paper, that in order to disorganize citizens, the government prefers certain forms of horizontal communication under its control.

Throughout this section, we assume $\rho_2 > \mu > \rho_1$.

4.1 Private Polling vs. Public Deliberation

The government can learn the citizens' preferences through private polling. In a private polling setup, citizens' messages can only be directly seen by the government. The following proposition shows that, private polling, in which the government is not constrained from manipulating citizens' learning, is always dominated either by public deliberation or no communication.

Proposition 7 ***(Public deliberation dominates private polling)*** *Providing $\rho_2 > \mu > \rho_1$, and λ is sufficiently small,*

(1) whenever the government finds it attractive to allow deliberation (i.e., public deliberation dominates no communication), it strictly prefers public deliberation to private polling; and

(2) whenever the government finds it unattractive to allow deliberation (i.e., no communication dominates public deliberation), it (weakly) prefers no communication to private polling.

Under private polling, the citizens can learn about each other's type by observing government policy changes. However, a fully separating equilibrium that is equivalent to the truth-telling equilibrium under public deliberation never exists,

because the government lacks a commitment device not to manipulate citizens' learning, and always has an incentive to deviate when both citizens file complaints.

As a result, there are only two types of possible equilibria under private polling. The first is a pooling equilibrium in which the government never adjusts policies and no information is revealed to citizens. The other is a semi-separating equilibrium in which government policy changes cannot reveal enough information for the citizens to perfectly infer each other's type. The government may choose the reform policy in response to two complaints with a probability smaller than one. Since an activist is not sure whether the other citizen is a non-activist when she sees the status quo policy, she may have a stronger incentive to join a protest than if she clearly knows that the other is a non-activist. The government thus loses a chance to disorganize the citizens when they are of different types since it cannot commit not to manipulate citizens' learning. Therefore, such an equilibrium gives the government strictly lower payoff than public deliberation. It implies that an authoritarian government can become better off by tying its hands in how it responds to public opinion so as to take full advantage of horizontal learning. The two mechanisms we propose in the next part further illustrate this idea. In the rest of this section, we assume the government prefers public deliberation to no communication.

4.2 Responsiveness Mechanism

The first mechanism we propose is called the *responsiveness* mechanism, in which the government is committed to the following rule:

> *When observing two complaints, the government implements the reform policy with probability σ; otherwise, it keeps the status quo policy.*

As is the case with private polling, in the responsiveness mechanism, the citizens cannot directly observe each other's messages. However, since the government is now able to choose a response rate (and commit to it) before it receives the citizens' complaints, the government can manipulate the citizens' belief of each other's type to its own advantage.

Since no one protests under the reform policy and a non-activist never protests (Lemma 1), we only need to pin down an activist's belief and strategy under the status quo policy. Given the status quo policy, the probability that an activist believes that the other citizen is also an activist is denoted by q. We can express q as a function of σ and γ, the preference correlation parameter that shapes the citizens' initial belief:

$$q \equiv \Pr(w_j = \overline{\omega} | w_i = \overline{\omega}, x = Q) = \frac{\gamma(1-\sigma)}{\gamma(1-\sigma) + (1-\gamma)}. \tag{18}$$

$q(\sigma)$ is strictly decreasing in σ. $q(0) = \gamma$, $q(1) = 0$, $q'(\sigma) = -\frac{\gamma(1-\gamma)}{(1-\gamma\sigma)^2} < 0$. By Lemma 1, an activist protests under the status quo with probability $p_0(q)$. Hence, the government's expected total payoff is:

$$E(\widetilde{G}) = -p\gamma[(1-\sigma)W(p_0(q)) + \sigma\mu] - 2p(1-\gamma)p_0(q(\sigma))\rho_1, \tag{19}$$

where $W(x) = \rho_2 x^2 + 2\rho_1 x(1-x)$.

When $\sigma = 1$, each citizen perfectly observes the other citizen's type. Such a mechanism represents the highest degree of horizontal openness and gives the government the same payoff as the truth-telling equilibrium under public deliberation. When $\sigma = 0$, horizontal communication is shut down and there is no policy change. Since the government policy is not conditional on complaints, vertical learning is

also of no use. Such a mechanism gives the government a weakly lower payoff than the pooling equilibrium when no deliberation is allowed.

Intuitively, when $\sigma \in (0,1)$, there must exist a certain form of horizontal learning in an indirect way. Yet, here we use a basic way to define whether the responsiveness mechanism exhibits horizontal learning. Whenever there are certian forms of horizontal learning, in most cases, it will be reflected in citizens' change of behaviors. Thus, we say *the responsiveness mechanism exhibits horizontal learning* if $p_0(q)$, the probability that an activist joins a protest under the status quo policy, is different from $p_0(\gamma)$, the probability that she does so with no communication.

We have already shown that public deliberation strictly dominates private polling, providing that the government prefers public deliberation to no communication. We also know that no communication gives the government a weakly higher payoff than $E(\widetilde{G})|_{\sigma=0}$. Therefore, as long as we can find a response rate $\widehat{\sigma}$ such that $E(\widetilde{G})|_{\widehat{\sigma}} > E(\widetilde{G})|_{\sigma=1}$, such a mechanism $\widehat{\sigma}$ is strictly preferred by the government to public deliberation, private polling and no deliberation. Suppose σ^* is the best among all the responsiveness mechanisms that we are looking at, i.e.,

$$\sigma^* \in \arg\max_{\sigma\in[0,1]} E(\widetilde{G}). \tag{20}$$

Proposition 8 shows that σ^* exists, and the response rate is moderate $\sigma^* \in (0,1)$, meaning that the moderate responsiveness mechanism is indeed preferred by the government to public deliberation and no communication. For convenience we refer to $q(\sigma^*)$ as q^*.

Proposition 8 ***(The moderate responsiveness mechanism)*** *Provided $\rho_2 > \mu > \rho_1$, the probability of success of an individual protest λ is sufficiently small, and the government prefers public deliberation to no communication,*

(1) the best responsiveness mechanism σ^* exists; (2) the optimal responsiveness is moderate, $\sigma^* \in (0,1)$; it gives the government a higher payoff than public deliberation, private polling and no communication; (3) $p_0(q^*) \in (p_0(0), p_0(\gamma))$ (the mechanism exhibits a certain form of horizontal learning); and (4) σ^* involves a commitment: under such a mechanism, the policy adjustment cost is strictly higher than the cost of collective action, i.e., $W(p_0(q^*)) < \mu$, thus the government has an ex post incentive to deviate.

Moderate responsiveness $\sigma^* \in (0,1)$, suggests that the government uses an optimal amount of horizontal communication to reshape the citizens' belief in its favored direction. As the level of horizontal openness increases, when an activist sees that the government keeps the status quo policy, she is more likely to believe that the other citizen is a non-activist, and thus has less incentive to join a protest. Under such a mechanism, although the government has to exert an extra amount of effort to implement the reform policy, it enjoys great opportunities to disorganize the citizens when they are of different types.

In order to make this work, it is necessary that the government is able to make a credible commitment to the citizens since it has an incentive to deviate from the rule when it receives two complaints. In the real world, one solution is to institutionalize the private and public deliberative process by setting up certain bureaucratic agencies (for example, the official petition system) to deal with citizens' complaints (Tsai and Xu, 2013; Lee and Zhang, 2013). The result of partial, but not full, responsiveness suggests that even if an authoritarian government is able to overcome the agency problems with local bureaucracies, it lacks the incentive to hold them fully accountable. This is consistent with the empirical findings that the official appeal system helps mediate social conflicts to some extent, but ponly

some popular claims are addressed by the authorities (Cai, 2004; Chen, 2012).

4.3 Censorship Mechanism

Alternatively, we could think of a strategy with which the government manipulates rhetoric in the public domain to disorganize the citizens. Consider the following mechanism:

> *When observing two complaints, the government reports a message "$\overline{\omega}$" to the public with probability ε and reports "$\underline{\omega}$" with probability $(1-\varepsilon)$; otherwise it reports "$\underline{\omega}$".*

Note that $\overline{\omega}$ and $\underline{\omega}$ are merely two labels. Again, in this mechanism, direct horizontal communication is not allowed; the government first elicits information from individuals privately, truncates the information, and sends it back to the citizens. We call it a *censorship* mechanism.

Such a mechanism can explain the online censorship of an authoritarian regime. The Chinese government hires Internet commentators to post favorable comments toward government policies as a way to sway public opinion.[26] Because of these maneuvers, the public discourse presented in front of the citizens is distorted. However, the government does not fully truncate "negative" news. King, Pan and Roberts (2013) recently report that the Chinese government allows citizens to criticize the government with much freedom, but actively censors information that arouses public anger or that can potentially spur collective action. Given the freedom of speech to a certain extent, even a sophisticated observer may believe that other citizens are likely to be satisfied with the government when she sees the

[26]The commentators are commonly called "fifty-cent party" by netizens as a satire since they are said to be paid RMB fifty cents for each post.

favorable comments about the government. Thus, she is discouraged from seriously challenging it. By garbling information in a limited way when public opinion is likely to be against the government, the government creates strategic uncertainties among citizens that help disorganize them.

Similarly as in Section 4.2, we write down the government's expected payoff under the censorship mechanism:

$$E(\widetilde{G}) = p\gamma[(1-\varepsilon)\max\{-W(p_0(q(\varepsilon))), -\mu\} - \varepsilon\mu] - 2p(1-\gamma)p_0(q(\varepsilon))\rho_1, \quad (21)$$

where $W(\cdot)$ is what is defined in Equation (19). Equation (21) differs from Equation (19) only in the term $\max\{-W(p_0(q(\varepsilon))) - \mu\}$. When two activists do not realize that they are both activists, under the censorship mechanism, the government does not need to commit to a policy change probabilistically as it does in the moderate responsiveness mechanism. The government always chooses the best policy following "sequential rationality." We formally define the optimal censorship mechanism ε^* similarly as:

$$\varepsilon^* \in \arg\max_{\varepsilon\in[0,1]} E(\widetilde{G}). \quad (22)$$

It turns out that at its optimum level the censorship strategy works in almost the same way as the best responsiveness mechanism. At their respective optima, the two mechanisms induce the same equilibrium outcome. Proposition 9 presents the results.

Proposition 9 ***(Equivalence of the two mechanisms)*** *Provided conditions in Proposition 8,*

(1) the best censorship mechanism ε^* exists; (2) it coincides with the best responsiveness mechanism σ^*, i.e., $\varepsilon^* = \sigma^*$, so that they induce the same equilibrium

outcome; and (3) when the two active citizens do not realize that they are of the same type, in the best censorship mechanism, the government keeps the status quo policy as a sequentially rational choice, i.e., $W(p_0(q(\varepsilon^*))) < \mu$.

The last point in the proposition indicates that the government is not committed to policy adjustment when it observes two complaints. The commitment is, however, at the censorship stage. In fact, the government has an incentive to deviate from the equilibrium strategy and disclose a less amount of "bad news" $\overline{\omega}$.

Loosely speaking, information matters because it helps the players distinguish some states from others. Hence, controlling information flows is not only about controlling the quantity of information, but also about the qualitative nature of information (Yang, 2013). In general, different forms of communication and different types of information control could lead to entirely different results for the government. By investigating the role of commitment in using horizontal learning, we illustrate that the government does not necessarily hate communication among the citizens, as long as it can design the form of communication appropriately. In fact, it prefers the kind of horizontal communication under its control that helps disorganize the citizens. Moreover, whether openness is preferred by the government also depends on information that is already dispersed in the society, that is, the common prior, based on which the citizens form their initial beliefs.

5 Conclusions

By introducing hierarchical communication into a collective-veto bargaining structure, this paper develops a tractable model to illustrate how an authoritarian state uses deliberative institutions to strengthen its rule. We emphasize that deliberation

is a process of hierarchical communication that incorporates both vertical communication between the citizens and the government, and horizontal communication among the citizens. The government can gain from vertical communication as it learns citizens' policy preferences and adjusts policies accordingly. Meanwhile, the government can either gain or lose from horizontal communication, depending on whether it impedes or fosters citizens' collective action by informing them of one another's preferences.

In general, a sophisticated authoritarian government allows citizens to complain in the public domain when the coordination effect from horizontal communication is dominated by the policy adjustment effect from vertical communication, and the disorganization effect from horizontal communication. The model also suggests that a deliberative authoritarian government could exhibit a certain degree of policy responsiveness provided that citizens' threat of collective action remains credible. In addition, it also shows that the relationship between the threat of collective action and regime openness is non-monotonic.

We emphasize that horizontal communication of certain forms is preferred by the government for the purpose of disorganizing citizens in collective action since it provides an opportunity for the government to reshape the citizens' beliefs. We show that as long as the government favors deliberation, it never wants to completely shut down horizontal communication even if it is able to do so. The government's best strategy is to manipulate citizens' horizontal communication, while also constraining itself from over-doing it, since manipulation compromises the credibility of the information.

To substantiate this idea, we investigate two institutional arrangements: the "moderate responsiveness mechanism," which combines private polling with a commitment to a moderate level responsiveness to citizen's complaints; and the "cen-

sorship mechanism," which truncates the information received from the citizens. We show that the two seemingly different mechanisms are essentially equivalent. To achieve the best payoff in both mechanisms, the government needs certain forms of commitment so as to utilize citizens' learning to create strategic uncertainties among them.

Our model captures some important aspects of the Chinese authoritarian rule. For example, in the absence of electoral competition or the rule of law, in many cases the government does listen to the citizens and respond very quickly on matters that many people care about and/or potential collective action is at stake (Nathan, 2003; Distelhorst, 2013). However, the top-down approach that the Chinese Communist Party (CCP) employs only tackles a small fraction of the citizens' appeals. This means a majority of the problems remains unheard or untouched by the government.

The model also explains, to a certain extent, why the information technology so far does not cause a huge problem for the authoritarian government. The rise of the Internet allows citizens to find information which was previously unaccessible to them, including one another's attitude towards the government policies. Nathan (2003) observes that the availability of the new technology is not likely to lead to a regime change in China since, quite paradoxically, routinized protests cannot send strong enough signals to trigger a large mass movement that is needed for a fundamental change. This paper offers an alternative explanation for the lack of a regime change. With effective online censorship, the government can not only learn from the citizens more efficiently and change policies that spur popular anger, but also use online deliberative platforms to disorganize the citizens.

However, the real world is certainly more complicated than the model. Harsh conditions drive desperate citizens to carry out more risky and high-profile ac-

tions against local officials to attract the attention of the media, the public and the higher-level government.[27] The seemingly endless stream of protests (mostly in the countryside) triggers the government's "maintaining social stability" (*weiwen*) strategy, which uses various kinds of semi-legal or illegal measures, including violence, to get control of the contentious state-society relations. Scholars argue that "maintaining social stability" strategy has significantly compromised the legitimacy of CCP's rule (Chen, 2012). This strategy changes the threat of citizens' collective action, thus jeopardizing government policy responsiveness. According to our model, its effect on regime openness is ambiguous.

[27] Their bold efforts are sometimes rewarded by the government, as Lorentzen (2013*a*) explains. Nevertheless this encourages even more contentious behaviors of the citizens.

References

Acemoglu, Daron and James Robinson. 2000. "Why did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics* 115(4):1167–1199.

Acemoglu, Daron and James Robinson. 2001. "A Theory of Political Transitions." *American Economic Review* 91(4):938–963.

Acharya, Avidit. 2012. "Coordination and Development in Dictatorships." Mimeo, University of Rochester.

Besley, Timothy and Robin Burgess. 2002. "The political economy of government responsiveness: Theory and evidence from India." *The Quarterly Journal of Economics* 117(4):1415–1451.

Boix, Carles and Milan Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions, Commitment, and Power-Sharing in Dictatorships." *The Journal of Politics* 75(02):300–316.

Botero, Juan, Alejandro Ponce and Andrei Shleifer. 2013. "Education, Complaints, and Accountability." Mimeo, Harvard University.

Bueno de Mesquita, Bruce and George W. Downs. 2005. "Development and Democracy." *Foreign Affairs* September–October(84):77–86.

Cai, Yongshun. 2004. "Managed Participation in China." *Political Science Quarterly* 119(3):425–451.

Chen, Xi. 2012. *Social Protest and Contentious Authoritarianism in China*. Cambridge: Cambridge University Press.

Distelhorst, Gregory. 2013. "Publicity-Driven Accountability in China." Doctoral Thesis, Massachusetts Institute of Technology.

Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *Review of Economic Studies* (forthcoming).

Egorov, Georgy and Konstantin Sonin. 2012. "Incumbency Advantage in Non-Democracies." Mimeo, Northwestern University.

Egorov, Georgy, Sergei Guriev and Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103(4):645–669.

Farrell, Joseph and Robert Gibbons. 1989. "Cheap Talk with Two Audiences." *The American Economic Review* 79(5):1214–1223.

Gehlbach, Scott and Konstantin Sonin. 2013. "Government Control of the Media." Mimeo, University of Wisconsin-Madison.

Gehlbach, Scott and Philip Keefer. 2011. "Investment without democracy: Ruling-party institutionalization and credible commitment in autocracies." *Journal of Comparative Economics* 39(2):123–139.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2011. "Democracy and Transparency." *The Journal of Politics* 73(04):1191–1205.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2013. "Transparency, Protest and Political (In)Stability." Mimeo.

King, Gary, Jennifer Pan and Margaret Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107(2):1–18.

Lee, Ching Kwan and Yonghong Zhang. 2013. "The Power of Instability: Unraveling the Microfoundations of Bargained Authoritarianism in China." *American Journal of Sociology* 118(6):1475–1508.

Little, Andrew T. 2012. "Elections, Fraud, and Election Monitoring in the Shadow of Revolution." *Quarterly Journal of Political Science* 7(3):249–283.

Little, Andrew T. 2013. "An Informational Theory of Elections." Mimeo, New York University.

Lohmann, Susanne. 1993. "A Signaling Model of Informative and Manipulative Political Action." *American Political Science Review* 87(2):319–333.

Lorentzen, Peter. 2013*a*. "Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime." *Quarterly Journal of Political Science* 8(2):127–158.

Lorentzen, Peter. 2013*b*. "Strategic Censorship." *American Journal of Political Science* (forthcoming).

Meirowitz, Adam, Massimo Morelli, Kristopher W. Ramsay and Francesco Squintani. 2013. "Mediation and Strategic Militarization." Mimeo, Princeton University.

Morris, Stephen and Hyun Song Shin. 2006. Heterogeneity and Uniqueness in Interaction Games. In *The Economy as an Evolving Complex System III: Current Perspectives and Future Directions*, ed. Lawrence E. Blume and Steven N. Durlauf. New York: Oxford University Press.

Nathan, Andrew. 2003. "Authoritarian Resilience." *Journal of Democracy* 14(1):6–17.

Palfrey, Thomas and Howard Rosenthal. 1985. "Voter Participation and Strategic Uncertainty." *The American Political Science Review* 79(1):62–78.

Passarelli, Francesco and Guido Tabellini. 2013. "Emotions and Political Unrest." IGIER working paper.

Shi, Tianjian. 1997. *Political Participation in Beijing*. Cambridge, Massachusetts: Harvard University Press.

Svolik, Milan. 2012. *The Politics of Authoritarian Rule*. Cambridge: Cambridge University Press.

Tsai, Lily and Yiqing Xu. 2013. "Outspoken Insiders: Who Complains About the Governmetnt in Authoritarian China?" Mimeo, Massachusetts Institute of Technology.

Yang, Ming. 2013. "Coordination with Flexible Information Acquisition." Mimeo, Duke University.

Appendix

The government's payoffs in equations

$$G_1(x,\omega_1,\omega_2)=\begin{cases} -\mu x & \text{if} \quad \omega_1=\omega_2=\underline{\omega} \\ -(1-x)\rho_2-\mu x & \text{if} \quad \omega_1=\omega_2=\overline{\omega} \\ -(1-x)\rho_1 p_0(0)-\mu x & \text{if} \quad \omega_1=\underline{\omega},\ \omega_2=\overline{\omega} \\ -(1-x)\rho_1 p_0(0)-\mu x & \text{if} \quad \omega_1=\overline{\omega},\ \omega_2=\underline{\omega} \end{cases}. \tag{A1}$$

$$G_0(x,\omega_1,\omega_2)=\begin{cases} -\mu x & \text{if} \quad \omega_1=\omega_2=\underline{\omega} \\ -(1-x)[\rho_2 p_0(\gamma)^2+\rho_1 2p_0(\gamma)(1-p_0(\gamma))]-\mu x & \text{if} \quad \omega_1=\omega_2=\overline{\omega} \\ -(1-x)\rho_1 p_0(\gamma)-\mu x & \text{if} \quad \omega_1=\underline{\omega},\ \omega_2=\overline{\omega} \\ -(1-x)\rho_1 p_0(\gamma)-\mu x & \text{if} \quad \omega_1=\overline{\omega},\ \omega_2=\underline{\omega} \end{cases}. \tag{A2}$$

The model

In the Appendix, we show most of the results in a more generalized and more abstract model. The model(s) in the main part of the paper can be seen as special example of it. First, we present the model.

$\Pr(\omega_i=\overline{\omega})=p, i=1,2$

$$\omega_j|_{\omega_i=\overline{\omega}}=\begin{cases} \overline{\omega} & \text{with probability} \quad \gamma \\ \underline{\omega} & \text{with probability} \quad 1-\gamma \end{cases}, \quad i\neq j, \tag{A3}$$

where $\gamma\in(\max\{0,1-\frac{1-p}{p}\},1)$, and

$$\omega_j|_{\omega_i=\underline{\omega}}=\begin{cases} \overline{\omega} & \text{with probability} \quad \frac{p}{1-p}(1-\gamma) \\ \underline{\omega} & \text{with probability} \quad 1-\frac{p}{1-p}(1-\gamma) \end{cases}, \quad i\neq j. \tag{A4}$$

The two citizens' *policy preferences* in the policy adjustment stage are $u_i(x,\omega_i,\omega_j), i=1,2$.

Collective-action payoff is characterized by

	participate	abstain
participate	$V_i^{11}(x,\omega_i,\omega_j)-k_i$	$V_i^{10}(x,\omega_i)-k_i$
abstain	$V_i^{01}(x,\omega_i,\omega_j)$	$V_i^{00}(x,\omega_i)$

where $k_i \tilde{} F(\cdot)[\underline{k},\overline{k}]$ is the private cost of participating in a collective action, which is private information, and only observed by citizen i after he observes government's policy x.

$$A_i(x,\omega_i,\omega_j) = V_i^{11}(x,\omega_i,\omega_j) - V_i^{01}(x,\omega_i,\omega_j) \tag{A5}$$

$$B_i(x,\omega_i) = V_i^{10}(x,\omega_i) - V_i^{00}(x,\omega_i) \tag{A6}$$

Assumptions

It can be verified that assumptions in the benchmark model in Section 2.1 and in the extended model in the Supplementary Appendix are special cases of the following assumptions.

Assumption 1 *$F(\cdot)$ is weakly concave; $f(k) = F'(k) > 0, \forall k \in [\underline{k}, \overline{k}]$.*

Assumption 2 *$\exists i \in I = \{1,2\}, A_i(R,\omega_i,\omega_j) \leq \underline{k}, B_i(R,\omega_i) \leq \underline{k}, B_{-i}(R,\omega_{-i}) \leq \underline{k}, \forall \omega_i,\omega_j,\omega_{-i} \in \{\underline{\omega},\overline{\omega}\}$; $\forall i \in I = \{1,2\}, A_i(Q,\omega_i = \underline{\omega},\omega_j) \leq \underline{k}, B_i(Q,\omega_i = \underline{\omega}) \leq \underline{k}, \forall \omega_j \in \{\underline{\omega},\overline{\omega}\}$.*

Assumption 3 *$A = A_i(R,\omega_i = \overline{\omega},\omega_j = \overline{\omega}), B = B_i(R,\omega_i = \overline{\omega})$ do not depend on i. $A > \overline{k} > B > \underline{k}$.*

Proofs

Lemma 2 *(Characterizing the collective-action stage) In any equilibrium when both citizens truthfully reveal their types when they are allowed to speak, we have:*
(1) when $x = R$, no one protests; and
(2) when $x = Q, \omega_i = \underline{\omega}$ never protests; $\omega_i = \overline{\omega}$ protests according to a cut-point strategy:

$$a_i(\overline{\omega}) = \begin{cases} 1 & \text{if} \quad k_i \leq k^* \\ 0 & \text{if} \quad k_i > k^* \end{cases} . \tag{A7}$$

If $\alpha = 0, k^* = T_0(\gamma)$; if $\alpha = 1, m_{-i} = \underline{\omega}, k^* = T_0(0)$; if $\alpha = 1, m_{-i} = \overline{\omega}, k^* = \overline{k}$. $T_0(\cdot)$ is uniquely and well defined by $T_0 = \max\{\gamma(A-B)F(T_0)+B, 1\}$, where $F(\cdot)$ is the distribution function of the collective action cost k_i. $T_0(\cdot)$ is strictly increasing in γ when $\gamma \leq \gamma_0 = \frac{1-B}{A-B}$; $T_0(\cdot) = 1$, when $\gamma \geq \gamma_0$. Define

$$p_0(\gamma) = F(T_0(\gamma)), \tag{A8}$$

which is the probability with which an activist protests.

Proof *of* **Lemma** 2

(a) Conditions $A_i(R,\omega_i,\omega_j) \leq \underline{k}, B_i(R,\omega_i) \leq \underline{k}$ in Assumption 2 suggest that it is a dominant strategy for citizen i not to protest under reform policy R. Expecting i's behavior, citizen j therefore also finds it profitable to abstain because $B_{-i}(R,\omega_{-i}) \leq \underline{k}$. As a result, under the reform policy, no one will protest.

(b) Under the status quo policy, conditions in Assumption 2, $A_i(Q,\omega_i=\underline{\omega},\omega_j) \leq \underline{k}$, $B_i(Q,\omega_i=\underline{\omega}) \leq \underline{k}$, imply that the non-activist type $\underline{\omega}$ will never protest. Thus the only uncertainty is to what extent the other citizen is an activist. We first investigate an activist's behavior without deliberation with the common knowledge γ. The other possibilities can be actually treated as special cases.

(b.1) We first claim that in equilibrium an activist will use a cut-point strategy

$$a_i(\overline{\omega}) = \begin{cases} 1 & \text{if} \quad k_i \leq k^* \\ 0 & \text{if} \quad k_i > k^* \end{cases} \tag{A9}$$

because his payoff gain in protest is:

$\gamma \Pr(j \text{ protest}|\omega_i=\overline{\omega})A + (1-\gamma \Pr(j \text{ protest}|\omega_i=\overline{\omega}))B - k_i$.

Now suppose i's cut-point is $k_i^*, i=1,2$.

(b.2) According to (b.1) the payoff gain of player i is therefore

$\gamma F(k_j^*)(A-B) + B - k_i$.

It then can be verified that the equilibrium condition is equivalent to

$$k_1^* = \min\{\gamma F(k_2^*)(A-B)+B, \overline{k}\}, \tag{A10}$$

$$k_2^* = \min\{\gamma F(k_1^*)(A-B)+B, \overline{k}\}. \tag{A11}$$

(b.3) When $\gamma \geq \frac{\overline{k}-B}{A-B}$, we claim that the unique solution to equation (A10) and equation (A11) is $k_1^* = k_2^* = \overline{k}$.

We can easily verify that $k_1^* = k_2^* = \overline{k}$ is a solution to the above equations and hence is an equilibrium. We need to check the other two possibilities.

Possibility 1: If at least one cut-point $k_i^* = \overline{k}$, then according to equation (A10) and equation (A11), the other cut-point automatically becomes the corner solution $\overline{k}$.

Possibility 2: Both cut-points are interior $k_1^*, k_2^* \in [B, \overline{k})$. Without loss of generality, let's assume $k_1^* \leq k_2^*$, so we get

$$\gamma F(k_2^*)(A-B)+B \leq \gamma F(k_1^*)(A-B)+B, \tag{A12}$$

therefore $k_1^* \geq k_2^*$ so that $k_1^* = k_2^* \in [B, \overline{k})$. Let's denote them as k^*, we then have

$$k^* = \gamma F(k^*)(A-B)+B. \tag{A13}$$

Because of Assumption 1, $\psi(x) \triangleq \gamma F(x)(A-B)+B-x$ is also weakly concave. In addition we have $\psi(\underline{k}) = B - \underline{k} > 0, \psi(\overline{k}) \geq 0$.

$\forall k \in (\underline{k}, \overline{k})$ can be represented by $k = \theta\underline{k} + (1-\theta)\overline{k}$ for some $\theta \in (0,1)$. So $\psi(k) \geq \theta\psi(\underline{k}) + (1-\theta)\psi(\overline{k}) > 0$. As a result, $k_1^* = k_2^* = \overline{k}$ is the unique equilibrium.

(b.4) When $0 < \gamma < \frac{\overline{k}-B}{A-B}$, we first claim that any equilibrium $k_1^*, k_2^* \in [B, \overline{k})$. If not, we must have

$$\overline{k} = \min\{\gamma F(k_i^*)(A-B) + B, \overline{k}\}, \tag{A14}$$

for some i. However the right hand side $\min\{\gamma F(k_i^*)(A-B)+B, \overline{k}\} = \gamma F(k_i^*)(A-B) + B < \overline{k}$ because $\gamma < \frac{\overline{k}-B}{A-B}$. So we get a contradiction, hence $k_1^*, k_2^* \in [B, \overline{k})$. Similarly as in (b.3), without loss of generality, let's assume $k_1^* \leq k_2^*$, so we get

$$\gamma F(k_2^*)(A-B) + B \leq \gamma F(k_1^*)(A-B) + B, \tag{A15}$$

therefore $k_1^* \geq k_2^*$ so that $k_1^* = k_2^* \in [B, \overline{k})$. Let's denote them as k^*, we then have

$$k^* = \gamma F(k^*)(A-B) + B. \tag{A16}$$

Because of Assumption 1 $\psi(x) \triangleq \gamma F(x)(A-B) + B - x$ is also weakly concave. In addition we have: $\psi(\underline{k}) = B - \underline{k} > 0, \psi(\overline{k}) = \gamma(A-B) - (\overline{k} - B) < 0$.

Because of continuity of $\psi(x)$, $\exists$ a solution $k^* \in (\underline{k}, \overline{k})$ such that $k^* = \gamma F(k^*)(A-B) + B$.

Because of concavity of $\psi(x)$, applying the same logic in (b.3), $\forall k \in (\underline{k}, k^*)$, $\psi(x) > 0$ and $\forall k \in (k^*, \overline{k})$, $\psi(x) < 0$. As a result, k^* is the unique cut-point equilibrium.

(b.4) Therefore without deliberation, the equilibrium of an activist is uniquely determined by the cut-point $k^* \triangleq T_0(\gamma)$, where $T_0(\gamma)$ is uniquely and well defined by

$$T_0(\gamma) = \min\{\gamma F(T_0(\gamma))(A-B) + B, \overline{k}\}. \tag{A17}$$

When $\gamma \geq \frac{\overline{k}-B}{A-B}, T_0(\gamma) = \overline{k}$. When $\gamma < \frac{\overline{k}-B}{A-B}$, $T_0(\gamma)$ is uniquely and well defined by equation (A16).

To rewrite the above equation, we get:

$$\gamma = \frac{T_0 - B}{(A-B)F(T_0)}. \tag{A18}$$

To show that $T_0(\gamma)$ is strictly increasing when $\gamma < \frac{\overline{k}-B}{A-B}$, we only need to show above well-defined function is strictly increasing in T_0 when $T_0 \geq B$. It is obvious that $\frac{T_0-B}{(A-B)F(T_0)}$ is differentiable, thus we have:

$$\frac{d\frac{T_0-B}{(A-B)F(T_0)}}{dT_0} = \frac{F(T_0) - (T_0 - B)f(T_0)}{(A-B)(F(T_0))^2}. \tag{A19}$$

We merely need to show $F(T_0) - (T_0 - B)f(T_0) > 0$.

Because $F(T_0)$ is differentiable, $\exists \xi \in [\underline{k}, T_0]$ s.t. $F(T_0) = F(\underline{k}) + f(\xi)(T_0 - \underline{k}) > f(\xi)(T_0 - B) \geq (T_0 - B)f(T_0)$.

The last inequality comes from concavity.

(c) When an activist faces an activist, their common knowledge is they are perfectly and positively correlated, so their incentives are exactly the same as in the case without deliberation with $\gamma = 1$. They will protest with probability 1.

(d) When an activist faces a non-activist, her payoff gain is just $B - k_i$ since a non-activist will never protest. Her cut-point is therefore B, which is equal to $T_0(0)$. ■

Lemma 3 *Provided conditions below, the truth-telling equilibrium exists:*

(I) sufficient conditions for incentive compatibility of a non-activist type $\underline{\omega}$:
$u_i(R, \underline{\omega}) \leq u_i(Q, \underline{\omega}), V_i^{00}(R, \underline{\omega}) \leq V_i^{01}(Q, \underline{\omega}, \overline{\omega}) \leq V_i^{00}(Q, \underline{\omega})$.

(II) sufficient conditions for incentive compatibility of an activist type $\overline{\omega}$:
$V_i^{00}(R, \overline{\omega}) \geq \max\{V_i^{10}(Q, \overline{\omega}), V_i^{00}(Q, \overline{\omega})\}, u_i(R, \overline{\omega}) \geq u_i(Q, \overline{\omega}), V_i^{00}(R, \overline{\omega}) \geq V_i^{11}(Q, \overline{\omega}) - E(k_i)$,
$V_i^{11}(Q, \overline{\omega}, \overline{\omega}) \geq V_i^{10}(Q, \overline{\omega}), V_i^{01}(Q, \overline{\omega}, \overline{\omega}) \geq V_i^{00}(Q, \overline{\omega})$.

Proof *of* **Lemma** 3

(a) First we claim that in any fully separating equilibrium, if q_t is the probability that the government will reform given that it observes t "complaint", i.e. t citizen(s) reveal that they are activists, we must have $q_2 \geq q_1 \geq q_0 = 0$.

If no one protests, then the cost without reform is 0. If there is only one activist, the cost without reform is $-\rho_1$ times some probability. When there are two activists, the cost without reform is $-\rho_2$ which is the largest. Because of the cost monotonicity we automatically have $q_2 \geq q_1 \geq q_0 = 0$.

(b) We check the payoff gain of the $\underline{\omega}$ type between claiming $\overline{\omega}$ and $\underline{\omega}$.

When $\omega_j = \underline{\omega}$, by claiming $\overline{\omega}$ she gets $q_1[u_i(R, \underline{\omega}) + V_i^{00}(R, \underline{\omega})] + (1 - q_1)[u_i(Q, \underline{\omega}) + V_i^{00}(Q, \underline{\omega})]$;

by claiming $\underline{\omega}$ she gets $[u_i(Q, \underline{\omega}) + V_i^{00}(Q, \underline{\omega})]$.

When $\omega_j = \overline{\omega}$,by claiming $\overline{\omega}$ she gets $q_2[u_i(R, \underline{\omega}) + V_i^{00}(R, \underline{\omega})] + (1 - q_2)[u_i(Q, \underline{\omega}) + V_i^{01}(Q, \underline{\omega}, \overline{\omega})]$;

by claiming $\underline{\omega}$ she gets $q_1[u_i(R, \underline{\omega}) + V_i^{00}(R, \underline{\omega})] + (1 - q_1)[u_i(Q, \underline{\omega}) + p_0(0)V_i^{01}(Q, \underline{\omega}, \overline{\omega}) + (1 - p_0(0))V_i^{00}(Q, \underline{\omega})]$.

When the conditions in (I) are satisfied and $q_2 \geq q_1$, we can easily check that the payoff gain is non-positive. Thus IC constraint is satisfied.

(c) We check the payoff gain of the $\overline{\omega}$ type between claiming $\overline{\omega}$ and $\underline{\omega}$.

When $\omega_j = \underline{\omega}$, by claiming $\overline{\omega}$ she gets

$q_1[u_i(R, \overline{\omega}) + V_i^{00}(R, \overline{\omega})] + (1 - q_1)[u_i(Q, \overline{\omega}) + E_k(\max\{V_i^{10}(Q, \overline{\omega}) - k_i, V_i^{00}(Q, \overline{\omega})\})]$;

by claiming $\underline{\omega}$ she gets $u_i(Q, \overline{\omega}) + E_k(\max\{V_i^{10}(Q, \overline{\omega}) - k_i, V_i^{00}(Q, \overline{\omega})\})$.

When $\omega_j = \overline{\omega}$,by claiming $\overline{\omega}$ she gets

$q_2[u_i(R, \overline{\omega}) + V_i^{00}(R, \overline{\omega})] + (1 - q_2)[u_i(Q, \overline{\omega}) + V_i^{11}(Q, \overline{\omega}, \overline{\omega}) - E_k(k_i)]$;

by claiming $\underline{\omega}$ she gets $q_1[u_i(R,\overline{\omega})+V_i^{00}(R,\overline{\omega})]+(1-q_1)[u_i(Q,\overline{\omega})+z]$, where $z=E_k(\max\{p_0(0)V_i^{11}(Q,\overline{\omega},\overline{\omega})+(1-p_0(0))V_i^{10}(Q,\overline{\omega})-k_i, p_0(0)V_i^{01}(Q,\overline{\omega},\overline{\omega})-(1-p_0(0))V_i^{00}(Q,\overline{\omega})\})$

When the conditions in (II) are satisfied and $q_2 \geq q_1$, we have

$u_i(R,\overline{\omega})+V_i^{00}(R,\overline{\omega}) \geq u_i(Q,\overline{\omega})+V_i^{11}(Q,\overline{\omega},\overline{\omega})-E_k(k_i)$

$V_i^{11}(Q,\overline{\omega},\overline{\omega})-E_k(k_i) \geq z$

The second inequality comes from the facts that:

$V_i^{11}(Q,\overline{\omega},\overline{\omega})-k_i \geq p_0(0)V_i^{11}(Q,\overline{\omega},\overline{\omega})+(1-p_0(0))V_i^{10}(Q,\overline{\omega})-k_i$

$\Longleftrightarrow V_i^{11}(Q,\overline{\omega},\overline{\omega}) \geq V_i^{10}(Q,\overline{\omega})$

$\&\ V_i^{11}(Q,\overline{\omega},\overline{\omega})-k_i \geq p_0(0)V_i^{01}(Q,\overline{\omega},\overline{\omega})-(1-p_0(0))V_i^{00}(Q,\overline{\omega})$

$\Longleftarrow V_i^{11}(Q,\overline{\omega},\overline{\omega})-k_i \geq V_i^{01}(Q,\overline{\omega},\overline{\omega})$

$\Longleftrightarrow A \geq k_i$

As a result, the payoff gain is non-negative. Thus IC constraint is satisfied. ■

Note 1 for Lemma 3

It can be verified that sufficient conditions in Lemma 3 are satisfied in the benchmark model of Section 2.1, and in the extended model in the Supplementary Appendix with the given parameter assumptions. As a result, Proposition 2 is supported by Lemma 3.

Note 2 for Lemma 3

Here we only provide sufficient conditions. Even when the conditions are violated, fully separating equilibrium may still exist. The basic idea of these conditions are: (1) An activist has incentives to induce her ideal policy for the policy payoff in the policy adjustment stage; (2) In the collective-action stage, an activist always wants to incentivize the other citizen to participate since the conditions imply a feature of free-riding or spillover effect in public good provision.

In terms of more interesting tradeoff in signaling, it is doable to relax the sufficient conditions and allow the possibility that: (1) an activist may gain relatively more in collective action than in the case when her ideal policy is implemented, and (2) a non-activist may pretend to be an activist to incentivize the other citizen and free ride.

Under the general prior p, we rewrite government's utility difference:

$$\min\{M,\mu\}-p\gamma\min\{\rho_2,\mu\}-2p(1-\gamma)\min\{p_0(0)\rho_1,\mu\}, \tag{A20}$$

where

$$M=p\gamma p_0(\gamma)^2\rho_2+2pp_0(\gamma)\rho_1(1-\gamma p_0(\gamma)). \tag{A21}$$

The following two lemmas will be in use.

Lemma 4 *Suppose $f : I \to f(I)$ is a real value function, I is an inertial on $\mathbb{R}$. $f(x)$ is concave and strictly increasing, then $f^{-1}(y)$ is convex.*

Proof *of* **Lemma** 4

$\forall y_1, y_2 \in f(I)$ and $y_1 < y_2, \forall \theta \in [0,1]$, suppose $x_1 = f^{-1}(y_1), x_2 = f^{-1}(y_2)$. We need to show

$f^{-1}(\theta y_1 + (1-\theta)y_2) \leq \theta f^{-1}(y_1) + (1-\theta)f^{-1}(y_2)$.

It is equivalent to $f[f^{-1}(\theta y_1 + (1-\theta)y_2)] \leq f[\theta f^{-1}(y_1) + (1-\theta)f^{-1}(y_2)]$

i.e., $\theta y_1 + (1-\theta)y_2 \leq f(\theta x_1 + (1-\theta)x_2)$

i.e., $\theta f(x_1) + (1-\theta)f(x_2) \leq f(\theta x_1 + (1-\theta)x_2)$.

It is exactly the concavity of $f(x)$. ■

Lemma 5 *Suppose $\frac{F^{-1}(y)-B}{y}$ is concave, then $p_0(\gamma)$ is convex in γ when $\gamma \leq \gamma_0$.*

Proof *of* **Lemma** 5

By Lemma 1, we know that $T_0(\gamma)$ is strictly increasing in γ when $\gamma \leq \gamma_0$, so $p_0(\gamma) = F(T_0(\gamma))$ is also strictly increasing in γ when $\gamma \leq \gamma_0$.

To apply Lemma 4, we only need to check that $p_0^{-1}(\cdot)$ is concave. When $\gamma \leq \gamma_0$, according to Lemma 1, $T_0(\gamma)$ is determined by $T_0(\gamma) = \gamma F(T_0(\gamma))(A-B) + B$

Thus p_0 is determined by:

$$F^{-1}(p_0) = \gamma p_0(A-B) + B, \tag{A22}$$

which is equivalent to $\gamma = \frac{F^{-1}(p_0)-B}{(A-B)p_0}$.

Because $\frac{F^{-1}(y)-B}{y}$ is concave by assumption, $p_0^{-1}(\cdot)$ is concave, therefore by Lemma 4, $p_0(\gamma)$ is convex in γ when $\gamma \leq \gamma_0$. ■

Proof *of* **Proposition** 3

First we provide the more generalized conditions which include the conditions in Proposition 3 as special cases.

$\rho_2 \geq 2\rho_1$,

$\rho_2 > \mu \geq \max\{F(B)\rho_1, 2pp_0(0)\rho_1\}$

$\mu > F(B)[2\rho_1 f(F(B)) + (\rho_2 - 2\rho_1)F(B) + 2\rho_1]$

(a) Given equation (A21),

$$\frac{dM}{d\gamma} = \zeta_1 p_0'(\gamma) + \zeta_2 \gamma 2p_0(\gamma)p_0'(\gamma) + \zeta_2 p_0(\gamma)^2, \tag{A23}$$

where $\zeta_1 = 2p\rho_1, \zeta_2 = p(\rho_2 - 2\rho_1)$

$$\frac{d^2M}{d\gamma^2} = \zeta_1 p_0''(\gamma) + 2\zeta_2(p_0(\gamma)p_0''(\gamma) + (p_0'(\gamma))^2)\gamma + 4\zeta_2 p_0(\gamma)p_0'(\gamma). \tag{A24}$$

So M is strictly increasing and convex in γ given we already know $p_0''(\gamma)$ is convex by Lemma 5, when $\gamma \leq \gamma_0$.

Thus the payoff gain function when $\gamma \leq \gamma_0$ is convex, and when $\gamma > \gamma_0$ is linear and therefore concave.

(b) payoff gain$|_{\gamma=0} = \min\{2pp_0(0)\rho_1, \mu\} - 2p\min\{p_0(0)\rho_1, \mu\} = 0$

payoff gain$|_{\gamma=1} = \min\{p\rho_2, \mu\} - p\min\{\rho_2, \mu\} > 0$.

(c)

$$\frac{d\text{payoff gain}}{d\gamma}|_{\gamma=0} = \frac{dM}{d\gamma}|_{\gamma=0} - p\mu + 2pp_0(0)\rho_1. \tag{A25}$$

According to equation (A23),

$$\frac{dM}{d\gamma}|_{\gamma=0} = \zeta_1 p_0'(0) + \zeta_2 p_0(0)^2. \tag{A26}$$

Recall that p_0 is determined by equation (A22), we have

$$\frac{1}{f(p_0(0))} p_0'(0) = p_0(0)(A - B). \tag{A27}$$

That is $p_0'(0) = f(p_0(0))p_0(0)(A-B)$. Put it into equation (A26) and equation (A26) we have

$$\frac{d\text{payoff gain}}{d\gamma}|_{\gamma=0} = p\{F(B)[2\rho_1 f(F(B))(A-B)+(\rho_2-2\rho_1)F(B)+2\rho_1]-\mu\}. \tag{A28}$$

Because $\mu > F(B)[2\rho_1 f(F(B)) + (\rho_2 - 2\rho_1)F(B) + 2\rho_1]$

we know that $\frac{d\text{payoffgain}}{d\gamma}|_{\gamma=0} > 0$.

(d) Since we have $\frac{d\text{payoffgain}}{d\gamma}|_{\gamma=0} > 0$, payoff gain$|_{\gamma=0} = 0$, payoff gain$|_{\gamma=1} => 0$,

result in (a) implies that the payoff gain is U-shaped and has only one zero point in $(0, 1)$. ■

Proposition 10 ***(Non-monotonic relationship between threat of collective action and openness)*** *$\exists\delta > 0$, such that $\forall\gamma \in [\gamma_0 - \delta, \gamma_0)$, for sufficiently small $\rho_1 > 0, \exists\overline{\rho} > \mu > \underline{\rho} > 0$, s.t. (1) the government's payoff difference between allowing and forbidding deliberation is V-shaped. The V-shaped function reaches the lowest kink at $\rho_2 = \mu$; and (2) the government chooses to open deliberation when the threat is very small or relatively large:*

$$\alpha^* = \begin{cases} 1 & \text{if } \rho_2 \in (\rho_1, \underline{\rho}) \cup (\overline{\rho}, +\infty) \\ 0 & \text{if } \rho_2 \in [\underline{\rho}, \overline{\rho}] \end{cases} . \tag{A29}$$

Proof *of* **Proposition** 10

First, we provide the more generalized conditions which include the conditions in Proposition 4 as special cases.

$\exists \delta > 0$ such that $\forall \gamma \in (\gamma_0 - \delta, \gamma_0)$, as long as
$\mu > \frac{2\rho_1}{(1-p_0(\gamma)^2)\gamma}[p_0(\gamma)(1-\gamma p_0(\gamma)) - (1-\gamma)p_0(0)]$,
$\mu > \rho_1$,
$1 - 2p + p\gamma > 0$,
ρ_2 has a non-monotonic effect.

Step (1)

Since $\mu > \rho_1$,when $\rho_2 < \mu$, we have $M < \mu$. Thus, the payoff gain in equation (A20) is first a decreasing function then an increasing function, and finally a constant, as ρ_2 becomes larger and larger. The first kink is μ. In the following, we only have to show that the function values at the left and the right are positive, and at μ is negative.

Step (2) As $\rho_2 \to +\infty$, the payoff gain $\geq \mu - p\gamma\mu - 2p(1-\gamma)\mu > 0$.

Step (3) As $\rho_2 \to \rho_1$, the payoff gain function converges to
$2p_0(\gamma)\rho_1 p - p\gamma p_0(\gamma)^2\rho_1 - p\gamma\rho_1 - p_0(0)\rho_1 2p(1-\gamma)$.

It is positive if and only if $2p_0(\gamma) - \gamma p_0(\gamma)^2 - \gamma - p_0(0)2(1-\gamma) > 0$

As $\gamma \to \gamma_0$, the left of the inequality becomes $2 - 2\gamma_0 - p_0(0)2(1-\gamma_0)$ which is always positive.

Step (4) As $\rho_2 \to \mu$, the payoff gain becomes
$p\gamma p_0(\gamma)^2\mu + 2p_0(\gamma)\rho_1 p(1-\gamma p_0(\gamma)) - p\gamma\mu - p_0(0)\rho_1 2p(1-\gamma)$,
which is negative given
$\mu > \frac{2\rho_1}{(1-p_0(\gamma)^2)\gamma}[p_0(\gamma)(1-\gamma p_0(\gamma)) - (1-\gamma)p_0(0)]$. ■

Proof *of* **Proposition** 7

(1) In any truth-telling equilibrium, because $\rho_2 > \mu > \rho_1$, government's equilibrium strategy is:

when observing $\omega_1 = \omega_2 = \overline{\omega}$, it implements reform with probability β^*;

in any other situations, it keeps the status quo policy.

(1.1)When $\beta^* = 1$, the equilibrium is exactly the same as the informative equilibrium in public deliberation, so that government is indifferent between private polling and public deliberation. However, when λ is small, $p_0(0) = F(B)$ as well as $W(p_0(0))$ is small. If the government deviates from the equilibrium strategy by keeping the status quo when facing two complaints, it will receive a strictly profitable payoff $-W(p_0(0)) > -\mu$. As a result, $\beta^* = 1$ can never be an equilibrium.

(1.2)When $\beta^* = 0$, the equilibrium is exactly the same as the pooling equilibrium without reform in public deliberation.

(1.3)When $\beta^* \in (0,1)$, when both citizens are activists, since the government is indifferent between reform and not reform (by the requirement of mixed strategy equilibrium), its payoff is $-\mu$ which is the same as the utility it gets with public deliberation.

Similarly, the payoff the government receives with two non-activists is also the same as in public deliberation.

When there is one activist and one non-activist, the utility is $-p_0(q)\rho_1$, where

$q = \Pr(w_j = \overline{\omega} | w_i = \overline{\omega}, \beta^*, x = Q) = \frac{\gamma(1-\beta^*)}{\gamma(1-\beta^*)+(1-\gamma)} > 0$.

Hence the payoff $-p_0(q)\rho_1$ is lower than the payoff with public deliberation $-p_0(0)\rho_1$ when the activist realizes for sure that the other citizen is a non-activist.

As a result, if $\beta^* \in (0,1)$, the government's payoff is strictly lower than its expected payoff with public deliberation.

(2.1) When public deliberation dominates no communication, case (1.3) directly implies that the equilibrium welfare of the government in public deliberation is higher than with private polling. Case (1.2) implies that no communication is weakly better than private polling. So public deliberation also dominates private polling.

(2.2) When no communication dominates public deliberation, with the similar logic, we can also show that the government's equilibrium welfare is higher with no communication than with private polling, which at most gives the government the highest welfare among public deliberation and no communication. ■

Proof *of* **Proposition** 8

Since $W(x) = (2\rho_2 - \rho_1)x^2 + 2\rho_1 x$, we get:

$$W(p_0(0)) = \rho_2 p_0(0)^2 + 2\rho_1 p_0(0)(1 - p_0(0)). \tag{A30}$$

(a) As long as $\rho_2 > \rho_1$,$W(x) = (2\rho_2 - \rho_1)x^2 + 2\rho_1 x$ is always strictly increasing when $x \in [0,1]$.

If $\rho_2 \geq 2\rho_1, W(x)$ is strictly increasing when $x > 0$;

If $\rho_2 < 2\rho_1, W(x)$ is strictly increasing when $x \in [0, \frac{\rho_1}{\rho_1 - (\rho_2 - \rho_1)}]$.

(b) Given the conditions provided in the proposition, the government's payoff under public deliberation $E(\widetilde{G})|_{\sigma=1}$ is greater than private polling and no communication. Hence, it is obvious that $\sigma^* > 0$. So we only need to show that $\sigma^* < 1$.

When λ is sufficiently small, we know that $B = \lambda L$ and $p_0(0) = F(B)$ are both sufficiently small. Therefore $p_0'(0) = f(p_0(0))p_0(0)(A - B)$ is also very small.

Thus $\mu > W(p_0(0)) + 2p_0'(0)\rho_1$, provided λ is sufficiently small. So the assumptions imply that

$$\max\{\rho_1, W(p_0(0)) + 2p_0'(0)\rho_1\} < \mu < \rho_2. \tag{A31}$$

(c) From equation (19) we get:

$$\frac{dE(\widetilde{G})}{d\sigma} = p\gamma[-(\mu - W(p_0(q))) - (1-\sigma)\frac{dW}{dp_0}\frac{dp_0}{dq}\frac{dq}{d\sigma}] - 2p(1-\gamma)\frac{dp_0}{dq}\frac{dq}{d\sigma}\rho_1, \tag{A32}$$

where $\frac{dW}{dp_0} > 0$, $\frac{dp_0}{dq} \geq 0$, $\frac{dq}{d\sigma} < 0$.

So we get

$$\frac{dE(\widetilde{G})}{d\sigma}|_{\sigma \to 1} = p\gamma[-(\mu - W(p_0(0))) + 2p_0'(0)\rho_1]. \tag{A33}$$

By inequality (A31), $\frac{dE(\widetilde{G})}{d\sigma}|_{\sigma\to 1} < 0$.

Because $E(\widetilde{G})$ is continuous in σ, σ^* always exists in [0,1]. Since $\frac{dE(\widetilde{G})}{d\sigma}|_{\sigma\to 1} < 0$,and $\sigma = 0$ is weakly dominated by $\sigma = 1$, we must have $\sigma^* \in (0,1)$.

(d) If $\gamma \leq \gamma_0$,we get $q^* \in (0,\gamma)$ so that $p_0(q^*) \in (p_0(0), p_0(\gamma))$ because $\sigma^* \in (0,1)$. If $\gamma > \gamma_0$, we still have $p_0(q^*) > p_0(0)$.

$$\frac{dE(\widetilde{G})}{d\sigma}|_{p_0(q)\geq 1} = p\gamma(\rho_2 - \mu) > 0. \tag{A34}$$

Suppose $\widehat{\sigma} = \sup\{\sigma : p_0(q(\sigma)) = 1\}$

$$\frac{dE(\widetilde{G})}{d\sigma}|_{\sigma\to\widehat{\sigma}^+} \geq p\gamma(\rho_2 - \mu) > 0 \tag{A35}$$

As a result, $\sigma^* > \widehat{\sigma}$ so that $p_0(q^*) < p_0(\gamma) = 1$.

Because $p_0(q^*) \in (p_0(0), p_0(\gamma))$ for $\forall\gamma \in (0,1)$, the best responsiveness mechanism involves a certain form of horizontal learning.

(e) The first order condition implies

$$\mu - W(p_0(q^*)) = -(1-\sigma^*)\frac{dW}{dp_0}\frac{dp_0}{dq}\frac{dq}{d\sigma} - 2\frac{1-\gamma}{\gamma}\frac{dp_0}{dq}\frac{dq}{d\sigma}\rho_1 \tag{A36}$$

so that $\mu - W(p_0(q^*)) > 0$. ■

Proof *of* **Proposition** *9*

We only need to show that ε^* in equation (22) is such that $W(p_0(q(\varepsilon^*))) < \mu$ so that the best censorship mechanism has the same maximizer as the best moderate responsiveness mechanism σ^*.

Because $W(p_0(q(\varepsilon)))$ is a decreasing function in $\varepsilon, \{\varepsilon : W(p_0(q(\varepsilon))) > \mu\}$ is either an empty set or $(0, \underline{\varepsilon})$. If the former is true, the optimization is exactly as in the responsiveness mechanism. If the latter is true,

$$\frac{dE(\widetilde{G})}{d\varepsilon}|_{\varepsilon\in(0,\underline{\varepsilon})} > 0. \tag{A37}$$

Thus the optimal ε^* always makes $W(p_0(q(\varepsilon^*))) < \mu$.

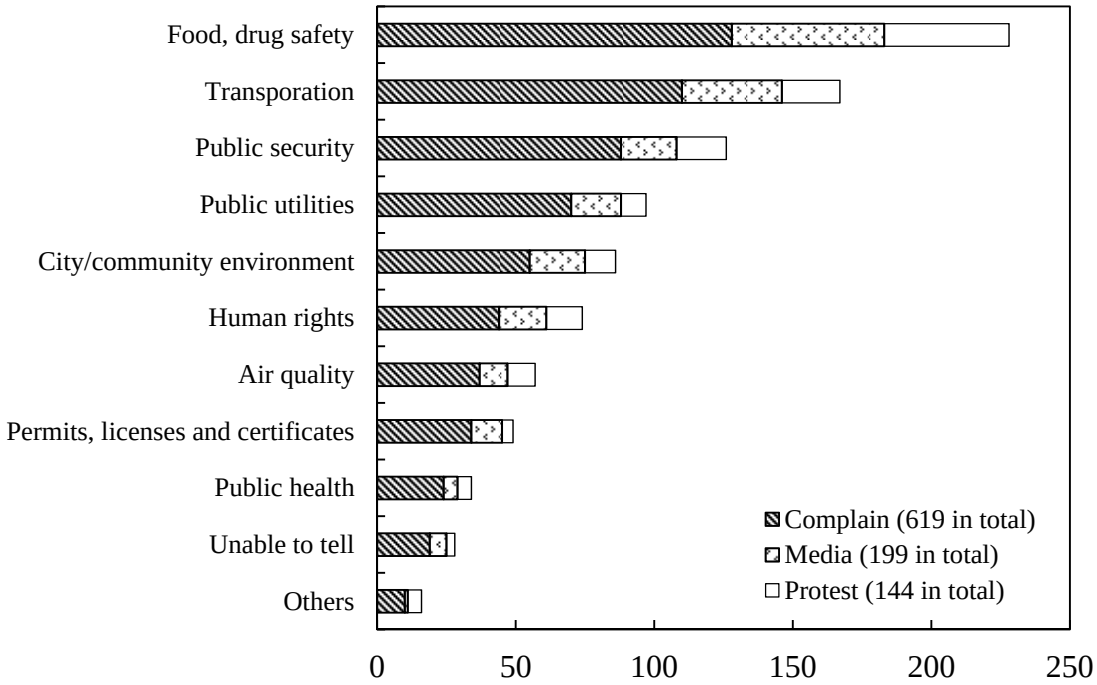
As a result, $\varepsilon^* = \sigma^*$, and these two mechanisms induce the same equilibrium outcome. ■

TABLE 1. FORMS OF POLITICAL ACTS

Forms of Political Acts	Count	Percentage
Complain to the government	**619**	**9.9%**
Go to government offices	241	3.9%
Call a government supervision hotline	218	3.5%
Use online measures, including government websites	197	3.1%
Go through local party committees/branches	129	2.1%
Contact local officials in private	129	2.1%
Collect signatures to send a petition	114	1.8%
Contact local officials directly	109	1.7%
Go for upper-level government	107	1.7%
Contact a representative of PC/PPCC	88	1.4%
Report to party disciplinary bodies	87	1.4%
Contact upper-level government officials in private	87	1.4%
Contact the media	**199**	**3.2%**
Collective action (small-scale protest)	**144**	**2.3%**
Total	**702**	**11.2%**

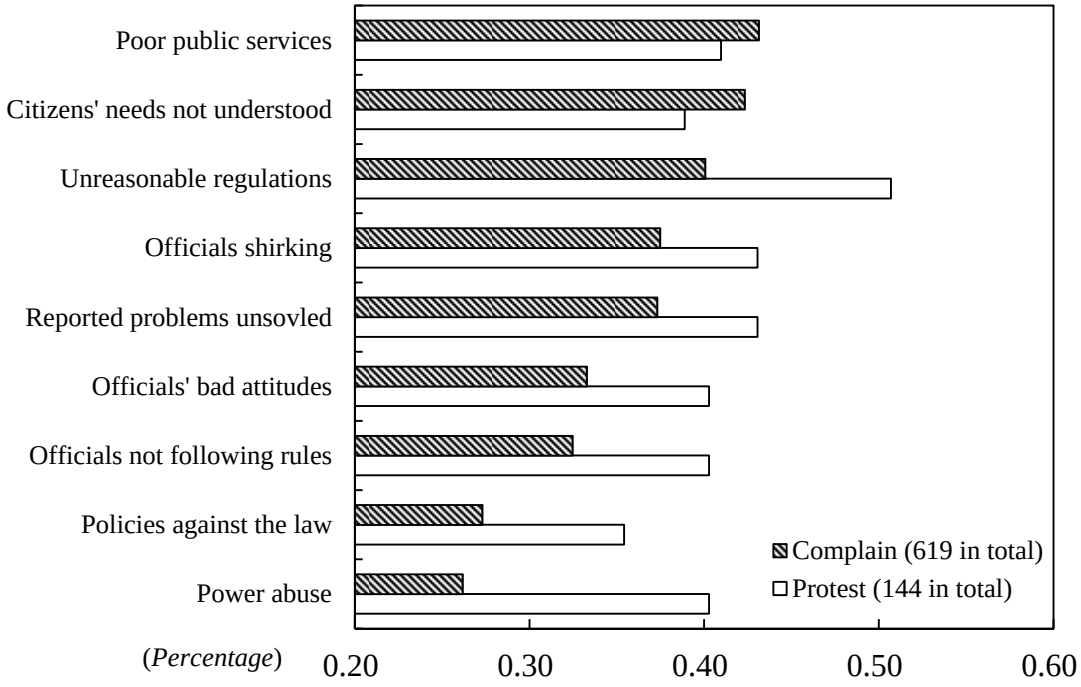
Note: The effective sample size is 6,259. The survey was conducted in 2013 in 30 large Chinese cities (all provincial capitals except Lhasa of Tibet). The respondents were asked: "During the past year, have you filed any complaints about the government and government officials in the following way?"

FIGURE 1. ISSUES OF CITIZENS' COMPLAINTS ABOUT THE GOVERNMENT



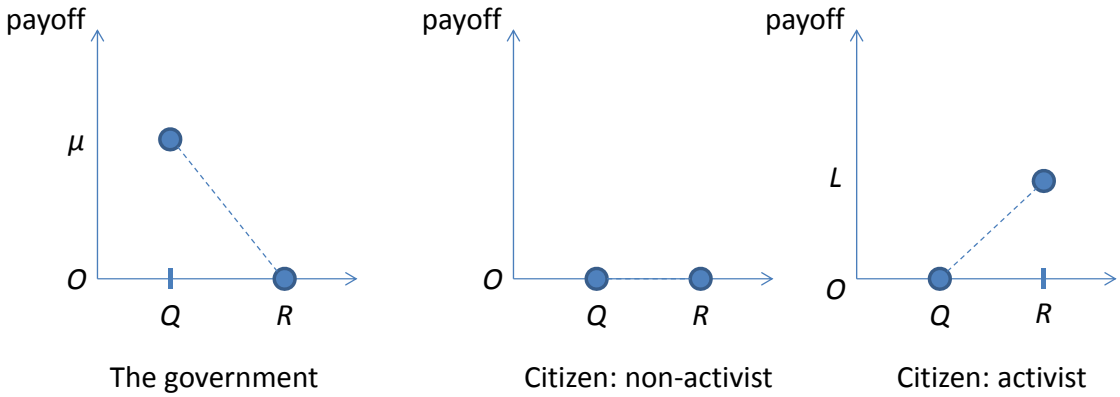
Note: The effective sample size is 6,259. The survey was conducted in 2013 in 30 large Chinese cities (all provincial capitals except Lhasa of Tibet). For the respondents who reported that they had complained about the local government or government official during the past year through government-provided channels (labeled "Complain"), the media (labeled "Media") and small-scale protest (label "Protest"), they were further asked the issues that their complaints were about.

FIGURE 2. WHAT ARE THE COMPLAINTS AND PROTESTS ABOUT?



Note: The effective sample size is 6,259. The survey was conducted in 2013 in 30 large Chinese cities (all provincial capitals except Lhasa of Tibet). For the respondents who reported that they had complained about the local government or government official during the past year through government-provided channels (labeled "Complain") or small-scale protest (label "Protest"), they were further asked why they were unsatisfied with the government.

FIGURE 3. PAYOFFS OF PLAYERS



The government | Citizen: non-activist | Citizen: activist

FIGURE 4. TIMELINE

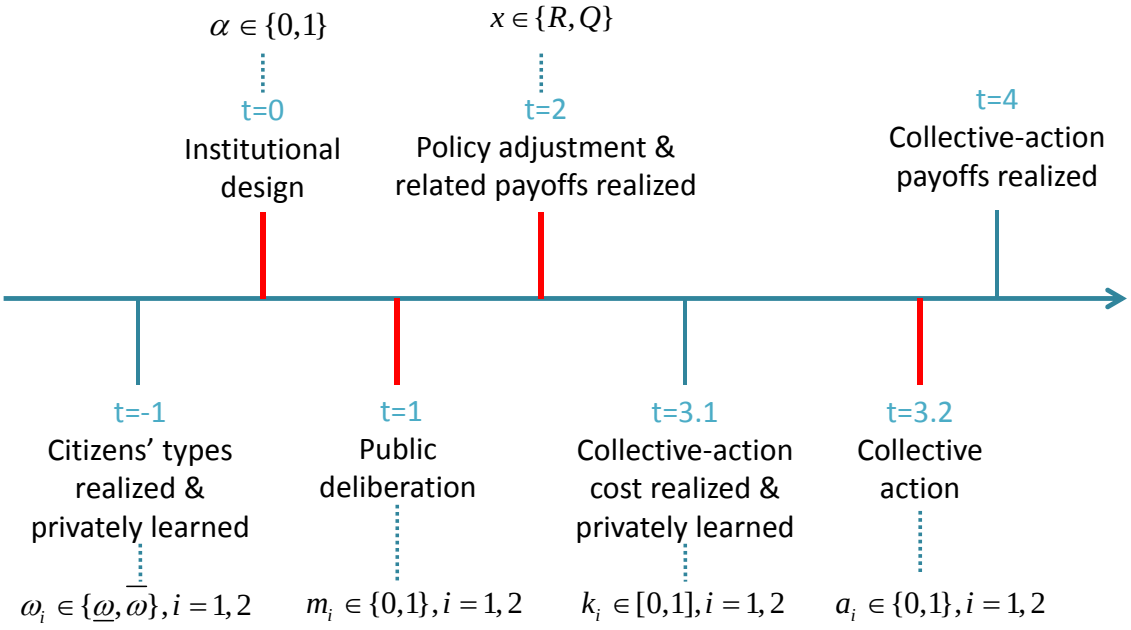


Figure 5. Probability of of an Activist Participating in Collective Action When Deliberation is Forbidden

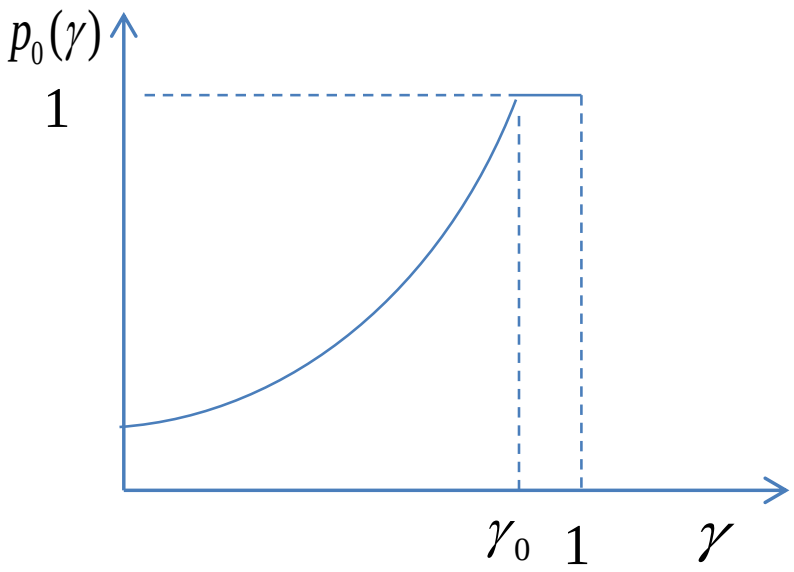


FIGURE 6. THE EFFECT OF HORIZONTAL COMMUNICATION ON PARTICIPATING IN COLLECTIVE ACTION

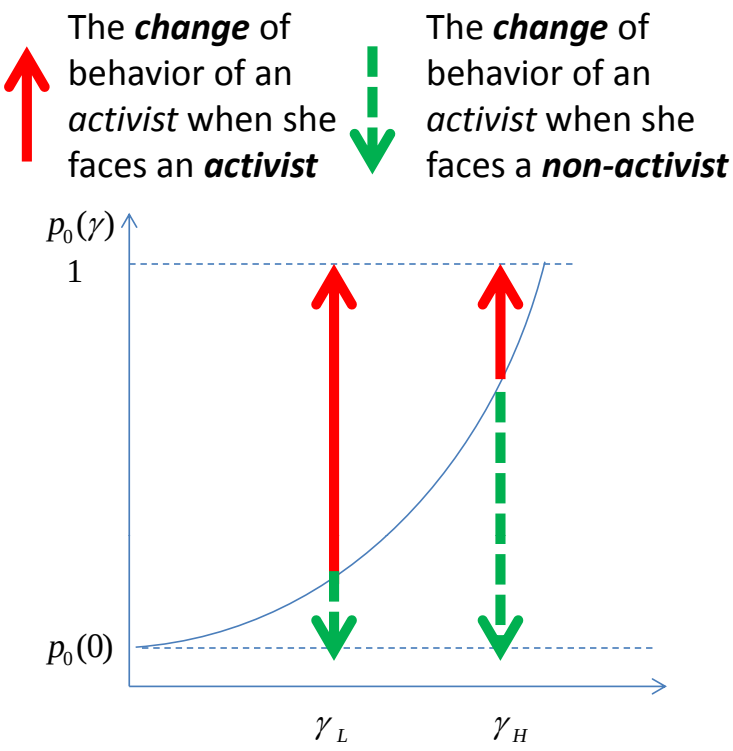


FIGURE 7. THE EFFECT OF CITIZENS' PREFERENCE CORRELATION ON THE GOVERNMENT'S CHOICE OF OPENING DELIBERATION

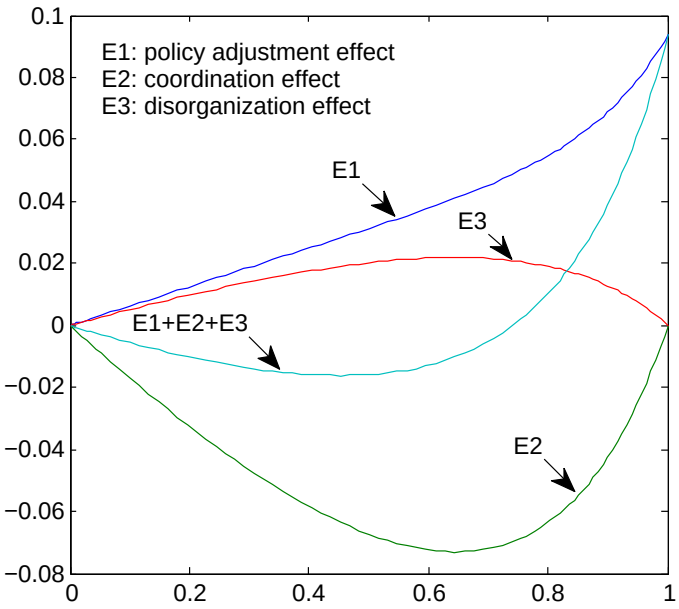


FIGURE 8. THE EFFECT OF COLLECTIVE-ACTION THREAT ON THE GOVERNMENT'S CHOICE OF OPENING DELIBERATION